WE CLAIM:

1. A genetic construct comprising a gene operatively-linked to a carrier, wherein the carrier is associated with a transmembrane form of viral glycoprotein or derivative thereof.

2. The genetic construct of Claim 1, wherein the transmembrane form of viral glycoprotein or derivative thereof is expressed on the surface of the carrier.

3. The genetic construct of Claim 1, wherein the transmembrane form of viral glycoprotein or derivative thereof is from Ebola.

4. The genetic construct of Claim 1, wherein the carrier is a viral vector.

5. The genetic construct of Claim 1, wherein the carrier is a non-biologic gene targeting vehicle.

6. The genetic construct of Claim 4, wherein the viral vector is a retroviral vector.

7. The genetic construct of Claim 4, wherein the viral vector is a lentiviral vector.

8. The genetic construct of Claim 5, wherein the non-biologic gene targeting vehicle is a liposome.

9. The genetic construct of Claim 5, wherein the non-biologic gene targeting vehicle is a DNA-protein complex.

10. A method of targeting a gene to a cell comprising the step of administering to a cell population a genetic construct comprising the gene operatively-linked to a carrier, wherein the carrier is associated with a transmembrane form of viral glycoprotein or derivatives thereof.

11. The method of Claim 10, wherein the transmembrane form of viral glycoprotein or derivative thereof is expressed on the surface of the carrier.

12. The method of Claim 10, wherein the transmembrane form of viral glycoprotein or derivative thereof is from Ebola.

13. The method of Claim 10, wherein the carrier is a viral vector.

14. The method of Claim 10, wherein the step of administration is *ex vivo*.

15. The method of Claim 10, wherein the step of administration is *in vivo*.

16. The method of Claim 10, wherein the cell is an endothelial cell.

17. The method of Claim 10, wherein the cell is a hepatocyte.

18. The method of Claim 10, wherein the cell is a monocyte.

19. The method of Claim 10, wherein the cell is a dendritic cell.

20. The method of Claim 14, further comprising the step of introducing the cell population to a subject.

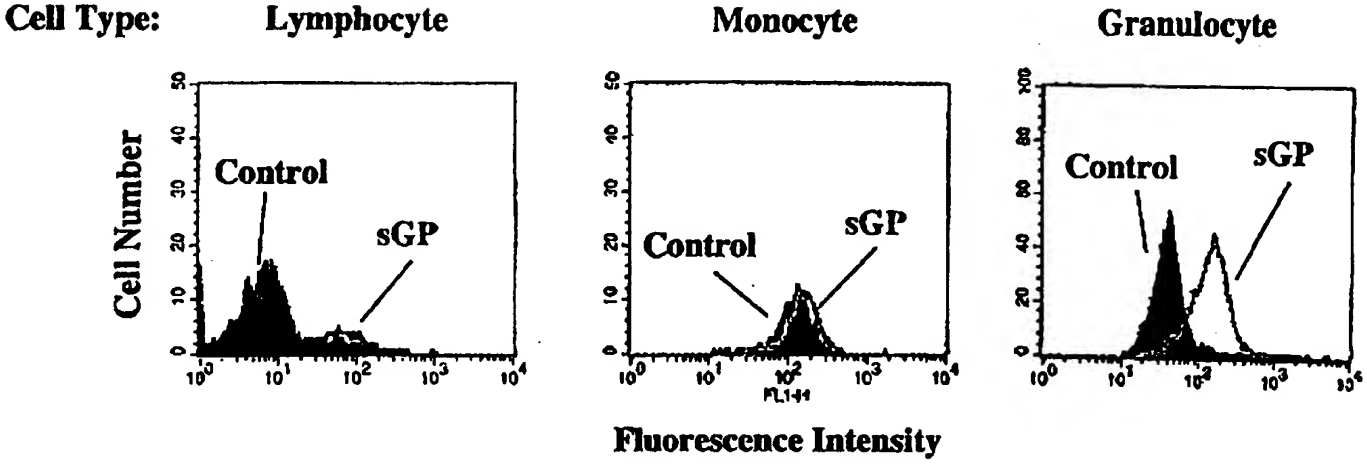


Figure 1A

Figure 1A1

Figure 1A2

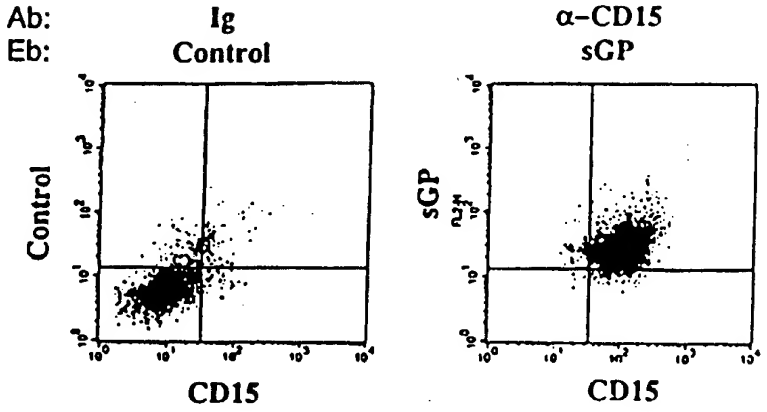


Figure 1B

Figure 1B1

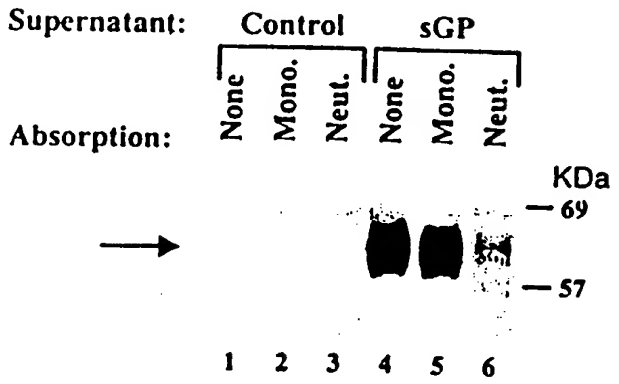


Figure 1C

WO 99/37331
1/11
PCT/US99/01382

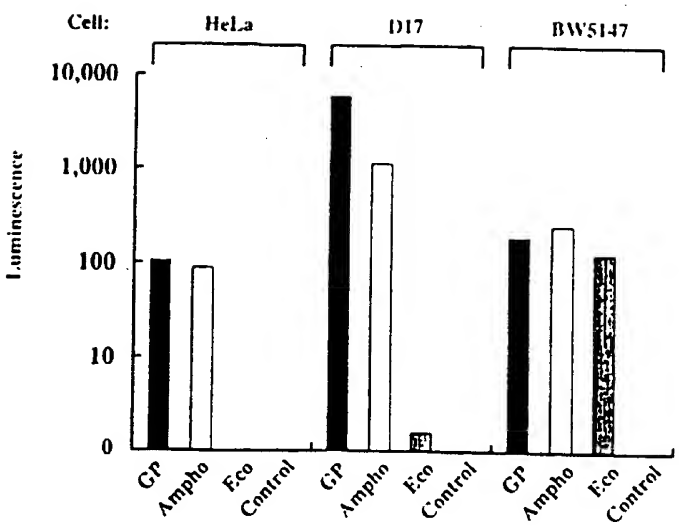


Figure 2A

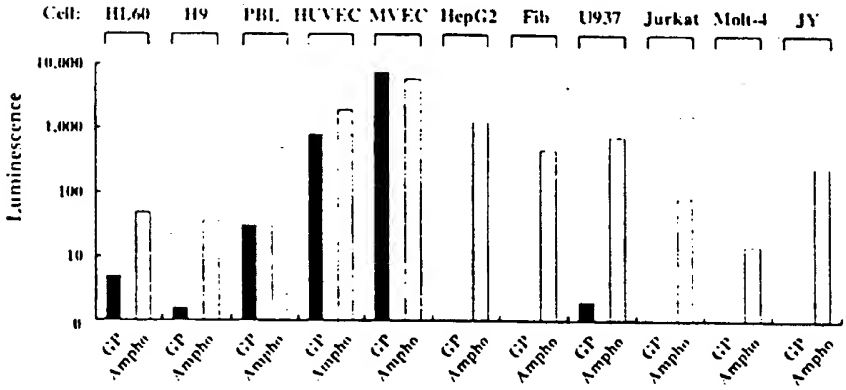


Figure 2B

Figure 2C

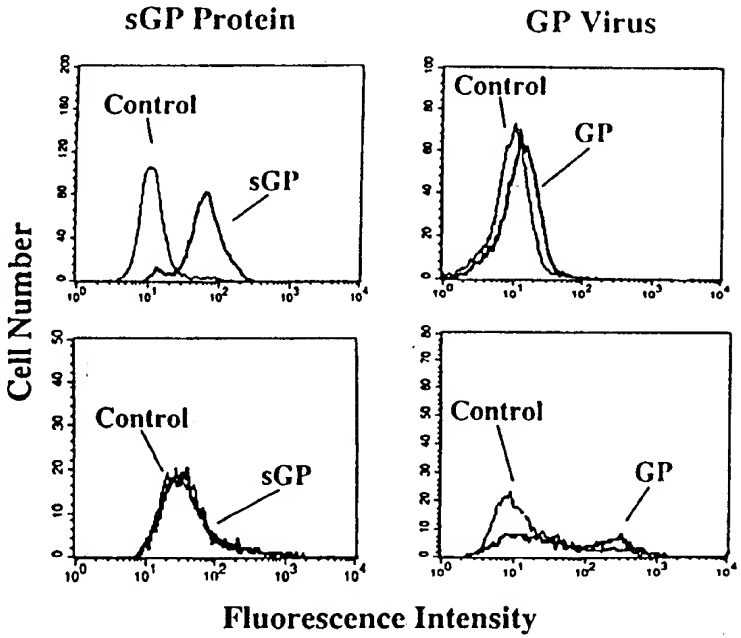


Figure 2C1

Figure 2C2

Figure 2C3

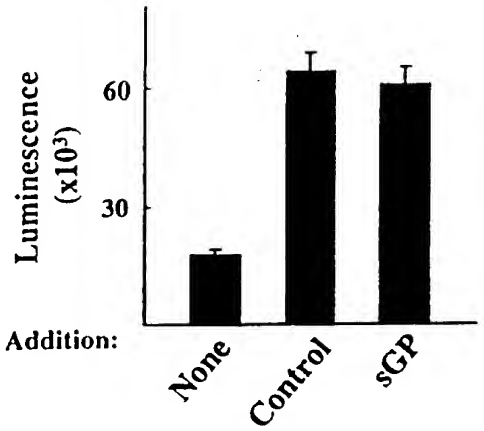


Figure 2D

Figure 3A

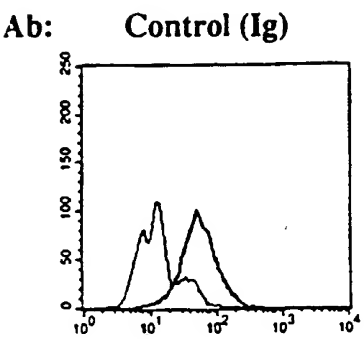


Figure 3B

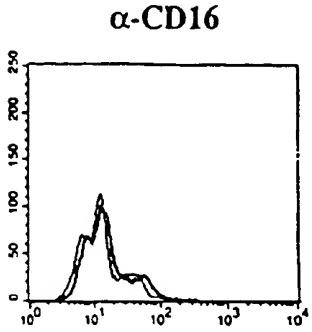


Figure 3C

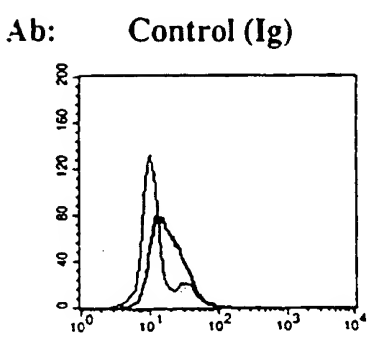


Figure 3D

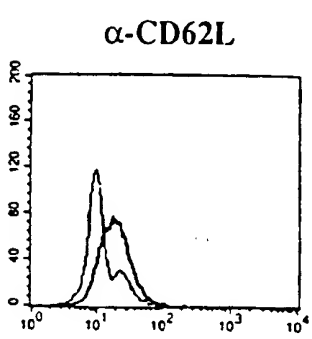


Figure 3E

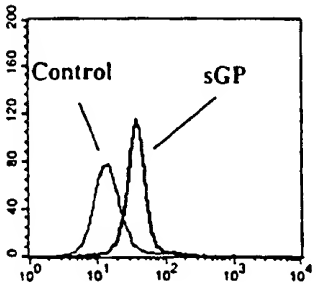
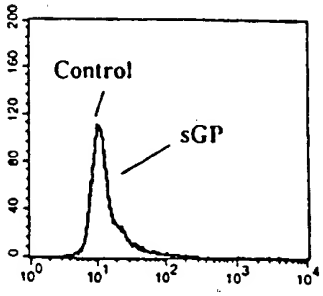


Figure 3F



Fluorescence Intensity

Figure 4A

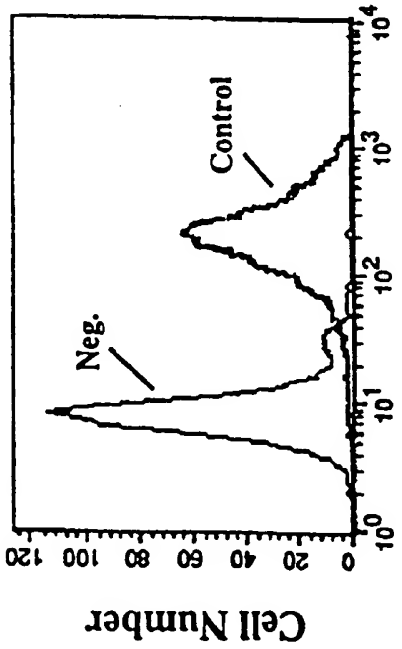
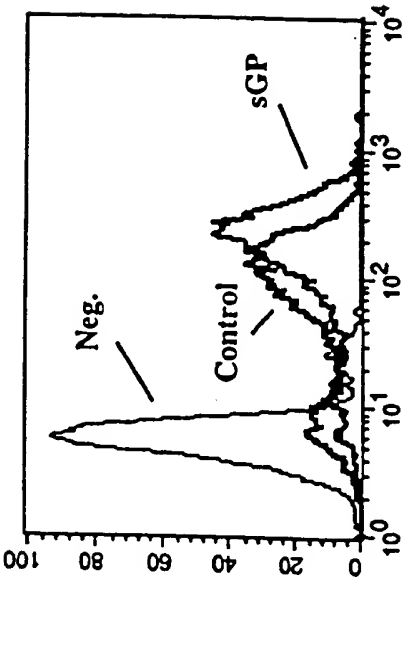


Figure 4B



Fluorescence Intensity

Figure 5A

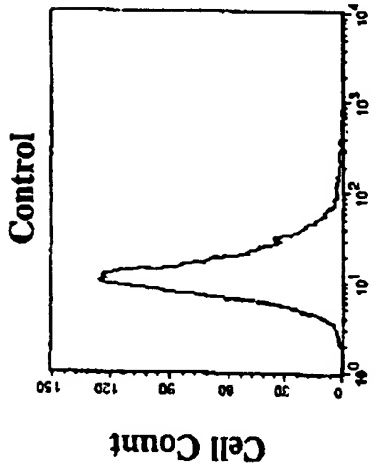


Figure 5B

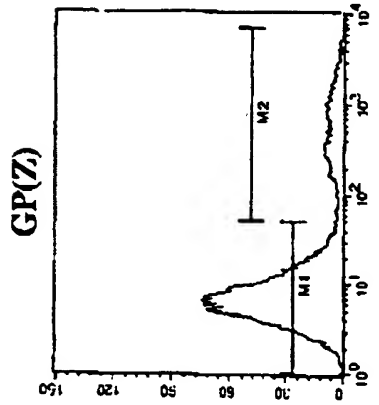
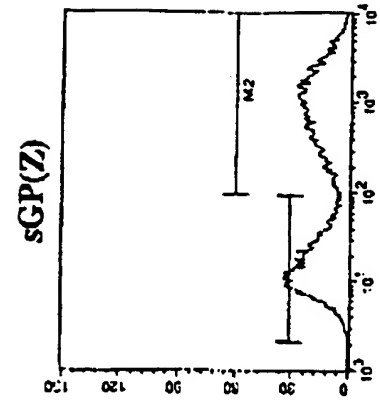


Figure 5C



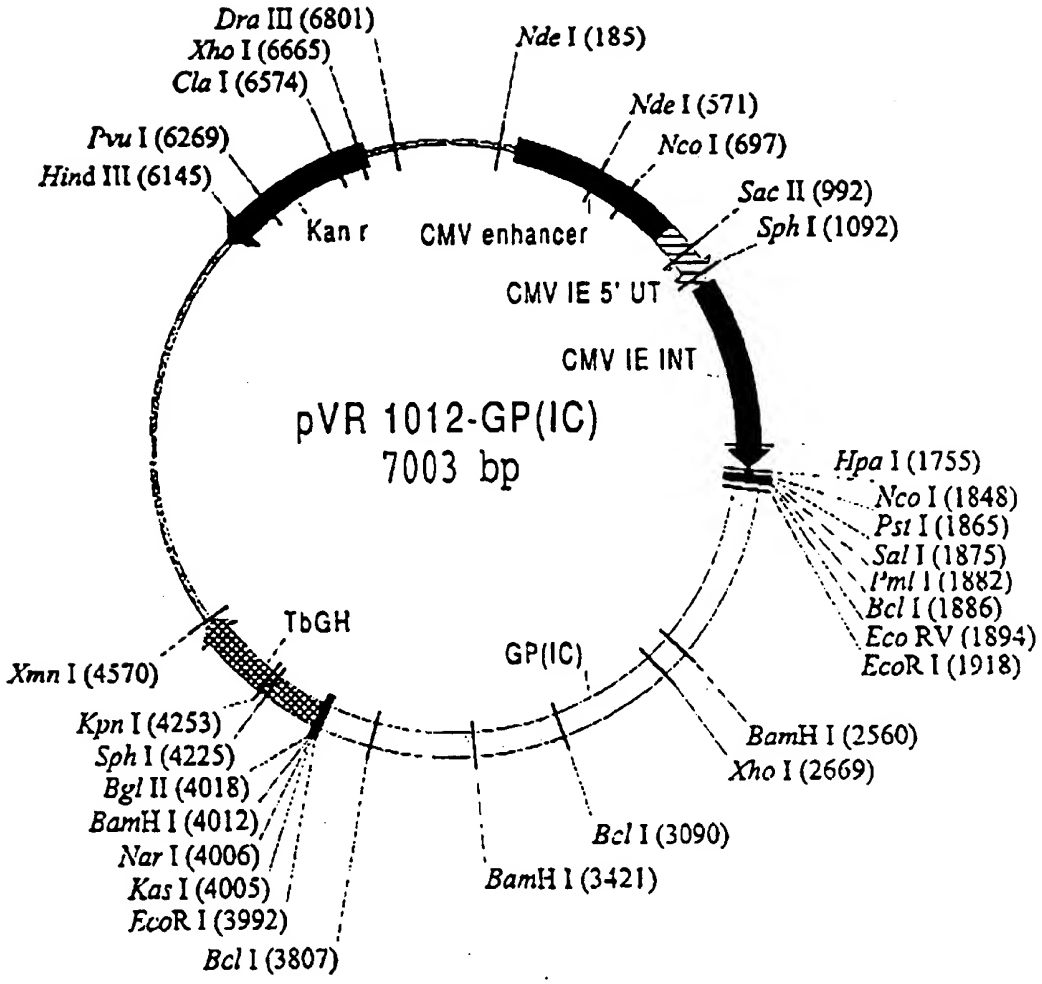


Figure 6

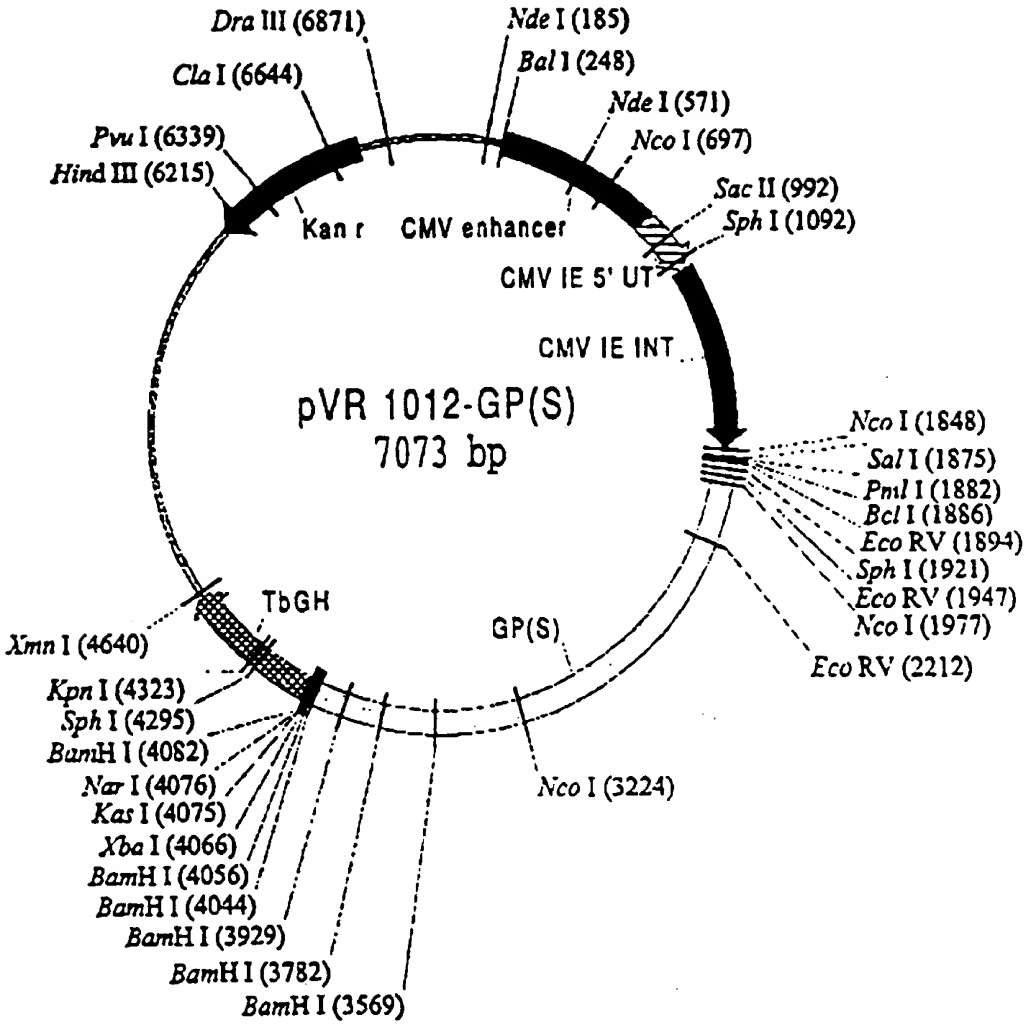

Dra III (6871)
Nde I (185)
Bal I (248)
Cla I (6644)
Nde I (571)
Pvu I (6339)
Nco I (697)
Hind III (6215)
Sac II (992)
Kan r
CMV enhancer
Sph I (1092)
CMV IE 5' UT
CMV IE INT
pVR 1012-GP(S)
7073 bp
Nco I (1848)
Sal I (1875)
Pml I (1882)
Bcl I (1886)
Eco RV (1894)
Sph I (1921)
Eco RV (1947)
Nco I (1977)
Eco RV (2212)
TbGH
Xmn I (4640)
GP(S)
Kpn I (4323)
Sph I (4295)
BamH I (4082)
Nar I (4076)
Kas I (4075)
Xba I (4066)
BamH I (4056)
BamH I (4044)
BamH I (3929)
BamH I (3782)
BamH I (3569)
Nco I (3224)

Figure 7

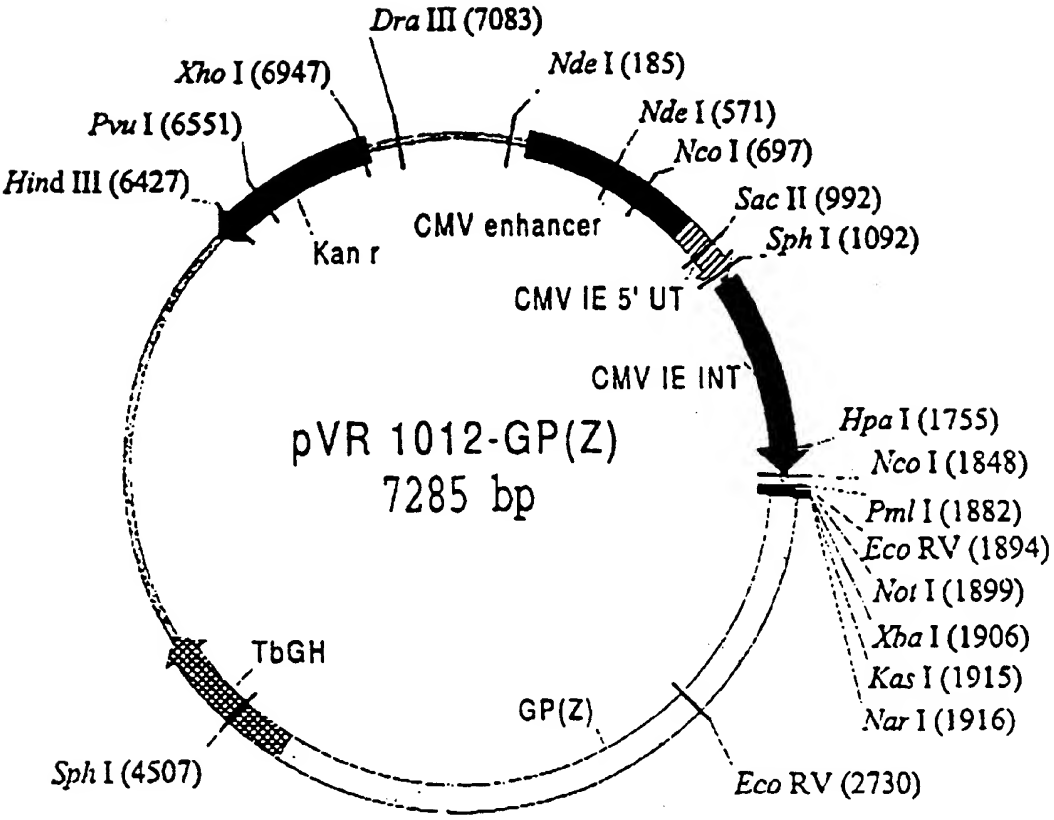


Figure 8

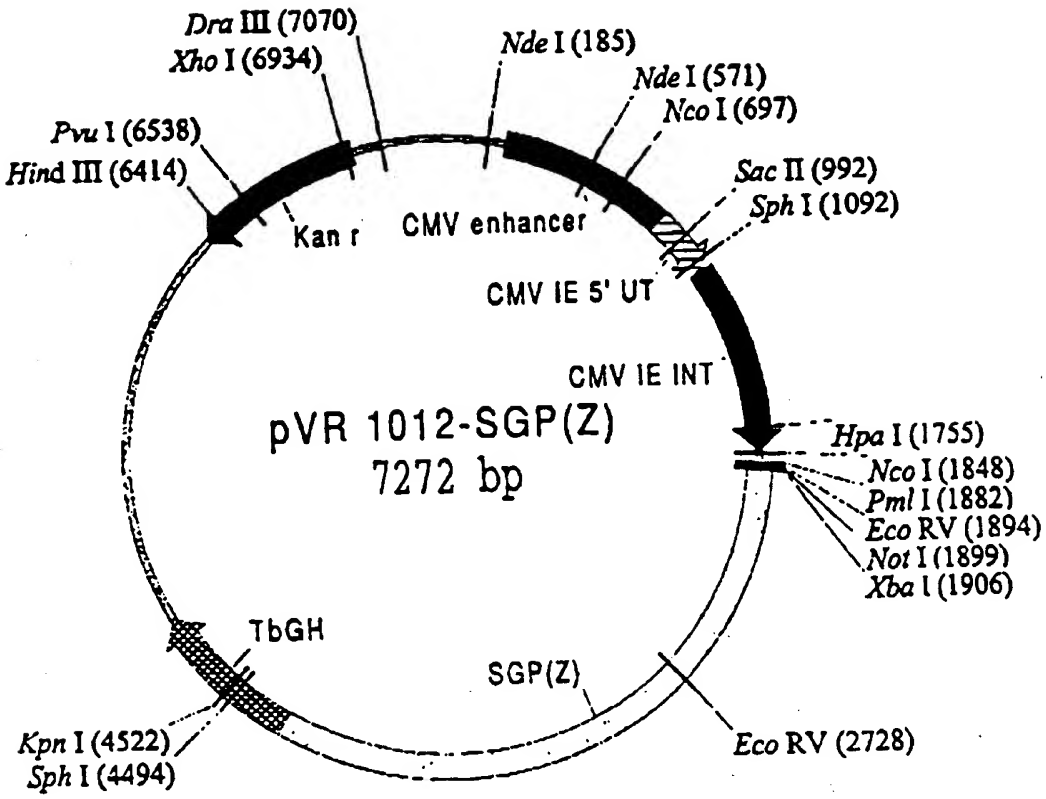

Dra III (7070)
Xho I (6934)
Nde I (185)
Nde I (571)
Nco I (697)
Pvu I (6538)
Hind III (6414)
Sac II (992)
Sph I (1092)
Kan r
CMV enhancer
CMV IE 5' UT
CMV IE INT
pVR 1012-SGP(Z)
7272 bp
Hpa I (1755)
Nco I (1848)
Pml I (1882)
Eco RV (1894)
Not I (1899)
Xba I (1906)
TbGH
SGP(Z)
Kpn I (4522)
Sph I (4494)
Eco RV (2728)

Figure 9

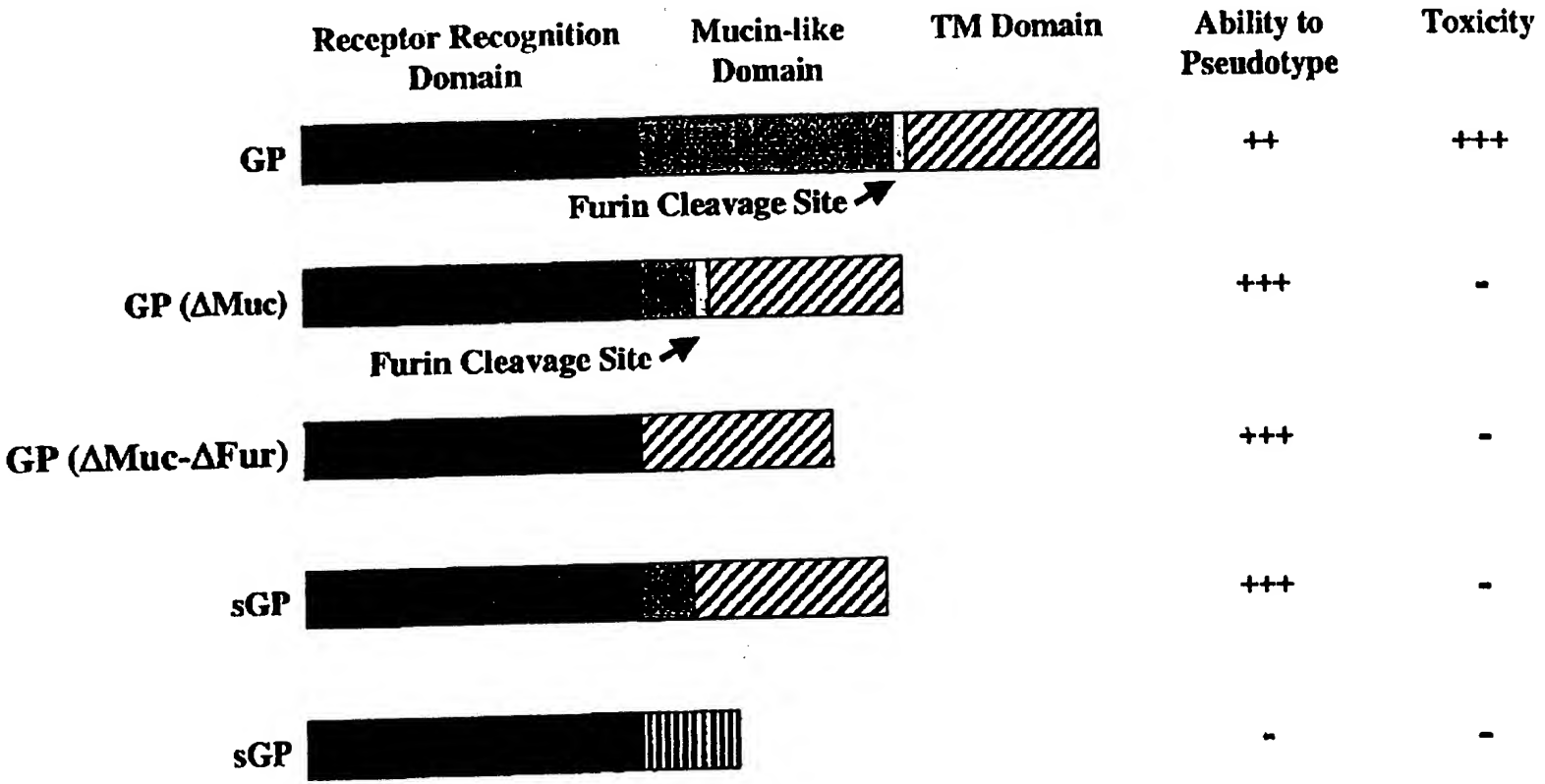

Receptor Recognition Domain
Mucin-like Domain
TM Domain
Ability to Pseudotype
Toxicity
GP
++
+++
Furin Cleavage Site
GP (ΔMuc)
+++
-
Furin Cleavage Site
GP (ΔMuc-ΔFur)
+++
-
sGP
+++
-
sGP
-
-

Figure 10

SEQUENCE LISTING ID NO: 1

pVR 1012-GP(IC)

General Description

DNA pVR 1012-GP(IC)
Local object
Created: 09/14/98 04:17PM
Last Modification Date: ? (no data)
length: 7003 bp
storage type: Basic
form: Circular

Comments

Restriction Map

BglII: 1 site AGATCT / TCTAGA

ClaI: 1 site ATCGAT / TAGCTA

DraIII: 1 site CACNNNGTG / GTGNNNCAC

EcoRV: 1 site GATATC / CTATAG

HindIII: 1 site AAGCTT / TTCGAA

HpaI: 1 site GTTAAC / CAATTG

KasI: 1 site GGCGCC / CCGCGG

KpnI: 1 site GGTACC / CCATGG

NarI: 1 site GGCGCC / CCGCGG

PmlI: 1 site CACGTG / GTGCAC

PstI: 1 site CTGCAG / GACGTC

PvuI: 1 site CGATCG / GCTAGC

SacII: 1 site CCGCGG / GGCGCC

SalI: 1 site GTCGAC / CAGCTG

XmnI: 1 site GAANNNNTTC / CTTNNNNAAG

EcoRI: 2 sites GAATTC / CTTAAG

NcoI: 2 sites CCATGG / GGTACC

NdeI: 2 sites CATATG / GTATAC

SphI: 2 sites GCATGC / CGTACG

XhoI: 2 sites CTCGAG / GAGCTC

BamHI: 3 sites GGATCC / CCTAGG

BclI: 3 sites TGATCA / ACTAGT

Functional Map

CDS (4 signals)

CMV IE 5' UT
Start: 886 End: 1129

CMV IE INT
Start: 1130 End: 1840

TbGH
Start: 4020 End: 4572

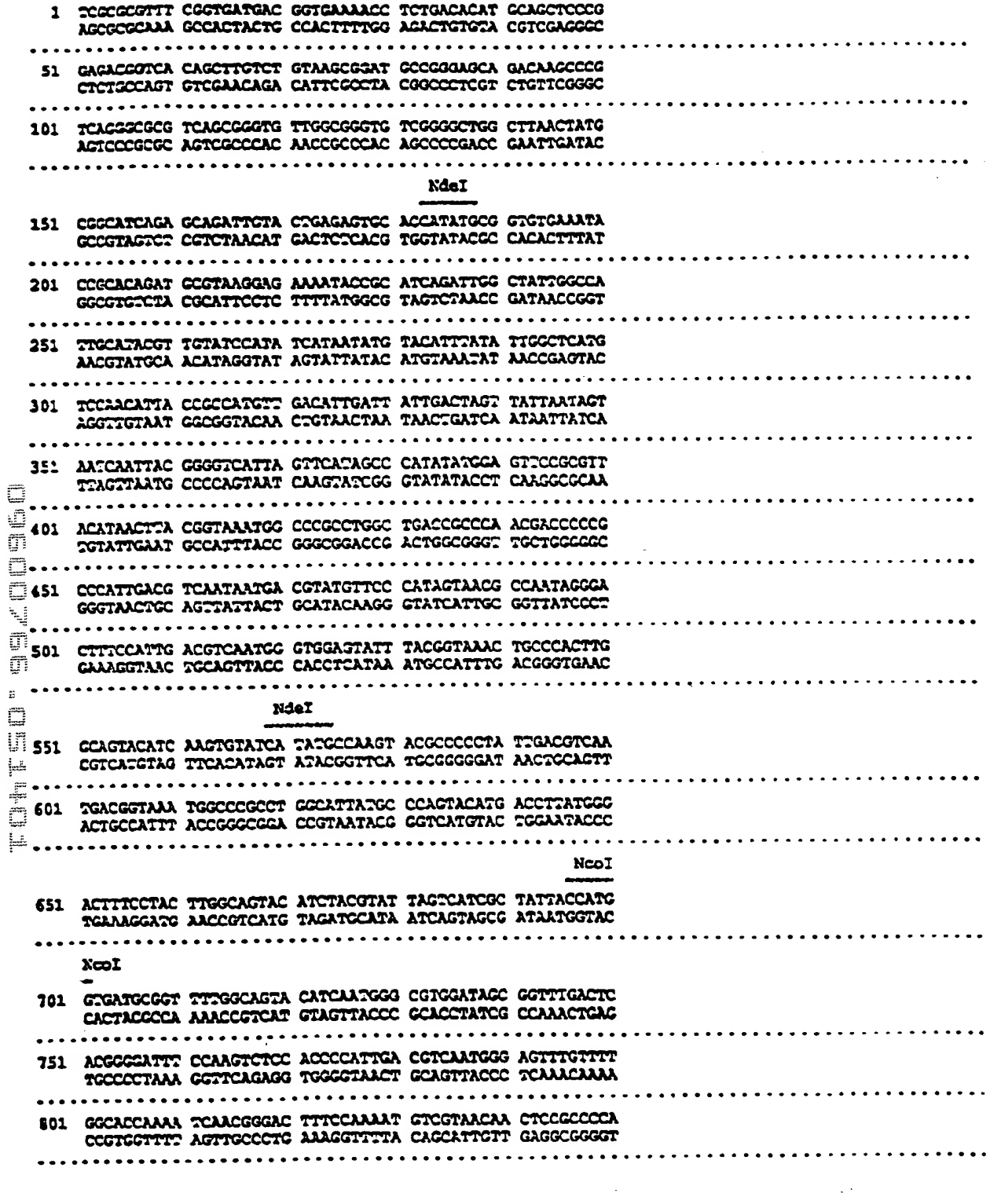
Kan r
Start: 6068 End: 6690 (Complementary)

Misc_feature (2 signals)

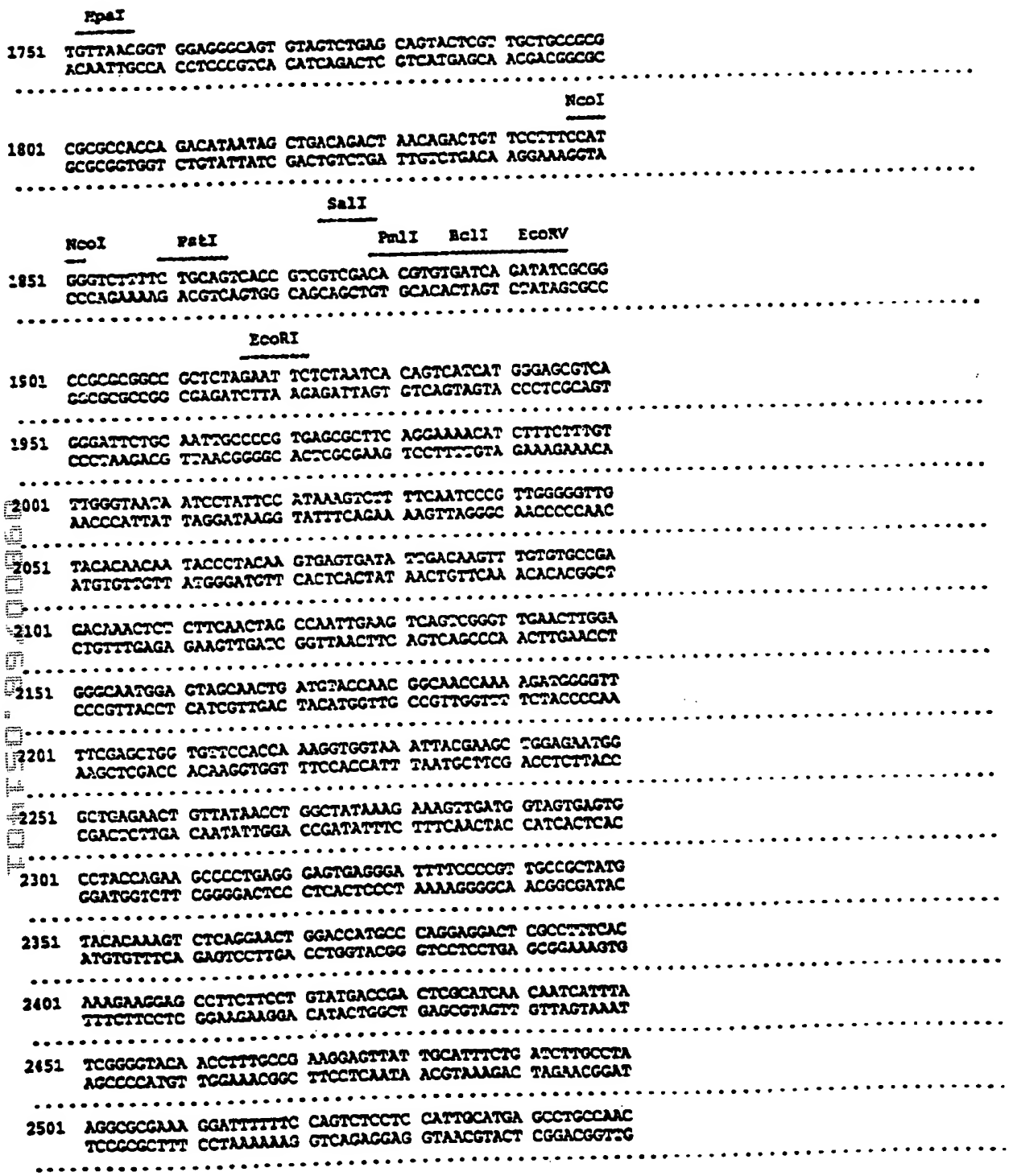
CMV enhancer
Start: 248 End: 885

GP(IC)
Start: 1870 End: 4019

Annotations


NdeI
NdeI
NcoI
NcoI

```
 851  TTGACGCAAA TGGGCGGTA  GCGTGTACGG TGGGAGGTCT ATATAAGCAG
      AACTGCGTTT ACCCGCCATC CGCACATGCC ACCCTCCAGA TATATTCGTC
 ....................................................................................

 901  AGCTCGTTTA GTGAACCGTC AGATCGCCTG GAGACGCCAT CCACGCTGTT
      TCGAGCAAAT CACTTGGCAG TCTAGCGGAC CTCTGCGGTA GGTGCGACAA
 ....................................................................................

                                                SacII
                                                ------

 951  TTGACCTCCA TAGAAGACAC CGGGACCGAT CCAGCCTCCG CGGCCGGGAA
      AACTGGAGGT ATCTTCTGTG GCCCTGGCTA GGTCGGAGGC GCCGGCCCTT
 ....................................................................................

1001  CGGTGCATTG GAACGCGGAT TCCCCGTCCC AAGAGTGACG TAAGTACCGC
      GCCACGTAAC CTTGCGCCTA AGGGGCAGGG TTCTCACTGC ATTCATGGCG
 ....................................................................................

                                                SphI
                                                ----

1051  CTATAGACTC TATAGGCACA CCCCTTTGGC TCTTATGCAT GCTATACTGT
      GATATCTGAG ATATCCGTGT GGGGAAACCG AGAATACGTA CGATATGACA
 ....................................................................................

1101  TTTTGGCTTG GGGCCTATAC ACCCCCGCTT CCTTATGCTA TAGGTGATGG
      AAAACCGAAC CCCGGATATG TGGGGGCGAA GGAATACGAT ATCCACTACC
 ....................................................................................

1151  CATAGCTTAG CCTATAGGTG TGGGTTATTG ACCATTATTG ACCACTCCCC
      ATATCGAATC GGATATCCAC ACCCAATAAC TGGTAATAAC TGGTGAGGGG
 ....................................................................................

1201  TATTGGTGAC GATACTTTCC ATTACTAATC CATAACATGG CTCTTTGCCA
      ATAACCACTG CTATGAAAGG TAATGATTAG GTATTGTACC GAGAAACGGT
 ....................................................................................

1251  CAACTATCTC TATTGGCTAT ATGCCAATAC TCTGTCCTTC AGAGACTGAC
      GTTGATAGAG ATAACCGATA TACGGTTATG AGACAGGAAG TCTCTGACTG
 ....................................................................................

1301  ACGGACTCTG TATTTTTACA GGATGGGGTC CCATTTATTA TTTACAAATT
      TGCCTGAGAC ATAAAAATGT CCTACCCCAG GGTAAATAAT AAATGTTTAA
 ....................................................................................

1351  CACATATACA ACAACGCCGT CCCCCGTGCC CGCAGTTTTT ATTAAACATA
      GTGTATATGT TGTTGCGGCA GGGGGCACGG GCGTCAAAAA TAATTTGTAT
 ....................................................................................

1401  GCGTGGGATC TCCACGCGAA TCTCGGGTAC GTGTTCCGGA CATGGGCTCT
      CGCACCCTAG AGGTGCGCTT AGAGCCCATG CACAAGGCCT GTACCCGAGA
 ....................................................................................

1451  TCTCCGGTAG CGGCGGAGCT TCCACATCCG AGCCCTGGTC CCATGCCTCC
      AGAGGCCATC GCCGCCTCGA AGGTGTAGGC TCGGGACCAG GGTACGGAGG
 ....................................................................................

1501  AGCGGCTCAT GGTCGCTCGG CAGCTCCTTG CTCCTAACAG TGGAGGCCAG
      TCGCCGAGTA CCAGCGAGCC GTCGAGGAAC GAGGATTGTC ACCTCCGGTC
 ....................................................................................

1551  ACTTAGGCAC AGCACAATGC CCACCACCAC CAGTGTGCCG CACAAGGCCG
      TGAATCCGTG TCGTGTTACG GGTGGTGGTG GTCACACGGC GTGTTCCGGC
 ....................................................................................

1601  TGGCGGTAGG GTATGTGTCT GAAAATGAGC GTGGAGATTG GGCTCGCACG
      ACCGCCATCC CATACACAGA CTTTTACTCG CACCTCTAAC CCGAGCGTGC
 ....................................................................................

1651  GCTGACGCAG ATGGAAGACT TAAGGCAGCG GCAGAAGAAG ATGCAGGCAG
      CGACTGCGTC TACCTTCTGA ATTCCGTCGC CGTCTTCTTC TACGTCCGTC
 ....................................................................................

1701  CCGAGTTGTT GTATTCTGAT AAGAGTCAGA GGTAACTCCC GTTGCGGTGC
      GACTCAACAA CATAAGACTA TTCTCAGTCT CCATTGAGGG CAACGCCACG
 ....................................................................................
```


FpaI
1751 TGTTAACGGT GGAGGCCAGT GTAGTCTGAG CAGTACTCGT TGCTGCCGCG
ACAATTGCCA CCTCCGGTCA CATCAGACTC GTCATGAGCA ACGACGGCGC
NcoI
1801 CGCGCCACCA GACATAATAG CTGACAGACT AACAGACTGT TCCTTTCCAT
GCGCGGTGGT CTGTATTATC GACTGTCTGA TTGTCTGACA AGGAAAGGTA
SalI
NcoI PstI PmlI BclI EcoRV
1851 GGGTCTTTTC TGCAGTCACC GTCGTCGACA CGTGTGATCA GATATCGCGG
CCCAGAAAAG ACGTCAGTGG CAGCAGCTGT GCACACTAGT CTATAGCGCC
EcoRI
1901 CCGCCCGGGC GCTCTAGAAT TCTCTAATCA CAGTCATCAT GGGAGCGTCA
GGCGCGCCCG CGAGATCTTA AGAGATTAGT GTCAGTAGTA CCCTCGCAGT
1951 GGGATTCTGC AATTGCCCCG TGAGCGCTTC AGGAAAACAT CTTTCTTTGT
CCCTAAGACG TTAACGGGGC ACTCGCGAAG TCCTTTTGTA GAAAGAAACA
2001 TTGGGTAATA ATCCTATTCC ATAAAGTCTT TTCAATCCCG TTGGGGGTTG
AACCCATTAT TAGGATAAGG TATTTCAGAA AAGTTAGGGC AACCCCCAAC
2051 TACACAACAA TACCCTACAA GTGAGTGATA TTGACAAGTT TGTGTGCCGA
ATGTGTTGTT ATGGGATGTT CACTCACTAT AACTGTTCAA ACACACGGCT
2101 GACAAACTCT CTTCAACTAG CCAATTGAAG TCAGTCGGGT TGAACTTGGA
CTGTTTGAGA GAAGTTGATC GGTTAACTTC AGTCAGCCCA ACTTGAACCT
2151 GGGCAATGGA GTAGCAACTG ATGTACCAAC GGCAACCAAA AGATCGGGTT
CCCGTTACCT CATCGTTGAC TACATGGTTG CCGTTGGTTT TCTAGCCCAA
2201 TTCGAGCTGG TGTTCCACCA AAGGTGGTAA ATTACGAAGC TGGAGAATGG
AAGCTCGACC ACAAGGTGGT TTCCACCATT TAATGCTTCG ACCTCTTACC
2251 GCTCAGAACT GTTATAACCT GGCTATAAAG AAAGTTGATG GTAGTGAGTG
CGACTCTTGA CAATATTGGA CCGATATTTC TTTCAACTAC CATCACTCAC
2301 CCTACCAGAA GCCCCTGAGG GAGTGAGGGA TTTTCCCCGT TGCCGCTATG
GGATGGTCTT CGGGGACTCC CTCACTCCCT AAAAGGGGCA ACGGCGATAC
2351 TACACAAAGT CTCAGGAACT GGACCATGCC CAGGAGGACT CGCCTTTCAC
ATGTGTTTCA GAGTCCTTGA CCTGGTACGG GTCCTCCTGA GCGGAAAGTG
2401 AAAGAAGGAG CCTTCTTCCT GTATGACCGA CTCGCATCAA CAATCATTTA
TTTCTTCCTC GGAAGAAGGA CATACTGGCT GAGCGTAGTT GTTAGTAAAT
2451 TCGGGGTACA ACCTTTGCCG AAGGAGTTAT TGCATTTCTG ATCTTGCCTA
AGCCCCATGT TGGAAACGGC TTCCTCAATA ACGTAAAGAC TAGAACGGAT
2501 AGGCGCGAAA GGATTTTTTC CAGTCTCCTC CATTGCATGA GCCTGCCAAC
TCCGCGCTTT CCTAAAAAAG GTCAGAGGAG GTAACGTACT CGGACGGTTG

BamHI

2551 ATGACCACGG ATCCCTCCAG TTACTATCAC ACGACAACAA TAAACTACGT
TACTGGTGCC TAGGGAGGTC AATGATAGTG TGCTGTTGTT ATTTGATGCA

2601 GGTTGATAAT TTTGGAACCA ACACCACAGA GTTTCTGTTC CAAGTCGATC
CCAACTATTA AAACCTTGGT TGTGGTGTCT CAAAGACAAG GTTCAGCTAG

XhoI

2651 ATTTGACGTA TGTGCAGCTC GAGGCAAGAT TCACACCACA ATTCCTTGTC
TAAACTGCAT ACACGTCGAG CTCCGTTCTA AGTGTGGTGT TAAGGAACAG

2701 CTCCTAAATG AAACCATCTA CCCTGATAAC CGCAGAAGTA ACACAACAGG
GAGGATTTAC TTTGGTAGAT GAGACTATTG GCGTCTTCAT TGTGTTGTCC

2751 AAAACTAATC TGGAAAATAA ATCCCACTGT TCATACCAGC ATGGGTGAGT
TTTTGATTAG ACCTTTTATT TAGGGTGACA AGTATGGTCG TACCCACTCA

2801 GGGCTTTCTG GGAAAATAAA AAAACTTCAC AAAAACCCTT TCAAGTGAAG
CCCGAAAGAC CCTTTTATTT TTTTGAAGTG TTTTTGGGAA AGTTCACTTC

2851 AGTTGTCTTT CGTACCTGTA CCAGAAACCC AGAACCAGGT CCTTGACACG
TCAACAGAAA GCATGGACAT GGTCTTTGGG TCTTGGTCCA GGAACTGTGC

2901 ACAGCGACGG TCTCTCCTCC CATCTCCGCC CACAACCACG CAGGCGAAGA
TGTCGCTGCC AGAGAGGAGG GTAGAGGCGG GTGTTGGTGC GTCCGCTTCT

2951 CCACAAAGAA TTGGTTTCAG AGGATTCCAC TCCAGTGGTT CAGATGCAAA
GGTGTTTCTT AACCAAAGTC TCCTAAGGTG AGGTCACCAA GTCTACGTTT

3001 ACATCAAGGG AAAGGACACA ATGCCAACCA CAGTGACGGG TGTACCAACA
TGTAGTTCCC TTTCCTGTGT TACGGTTGGT GTCACTGCCC ACATGGTTGT

BclI

3051 ACCACACCCT CTCCATTTCC AATCAATGCT CGCAACACTG ATCATACCAA
TGGTGTGGGA GAGGTAAAGG TTAGTTACGA GCGTTGTGAC TAGTATGGTT

3101 ATCATTTATC GGCCTGGAGG GGCCCCAAGA AGACCACAGC ACCACACAGC
TAGTAAATAG CCGGACCTCC CCGGGGTTCT TCTGGTGTCG TGGTGTGTCG

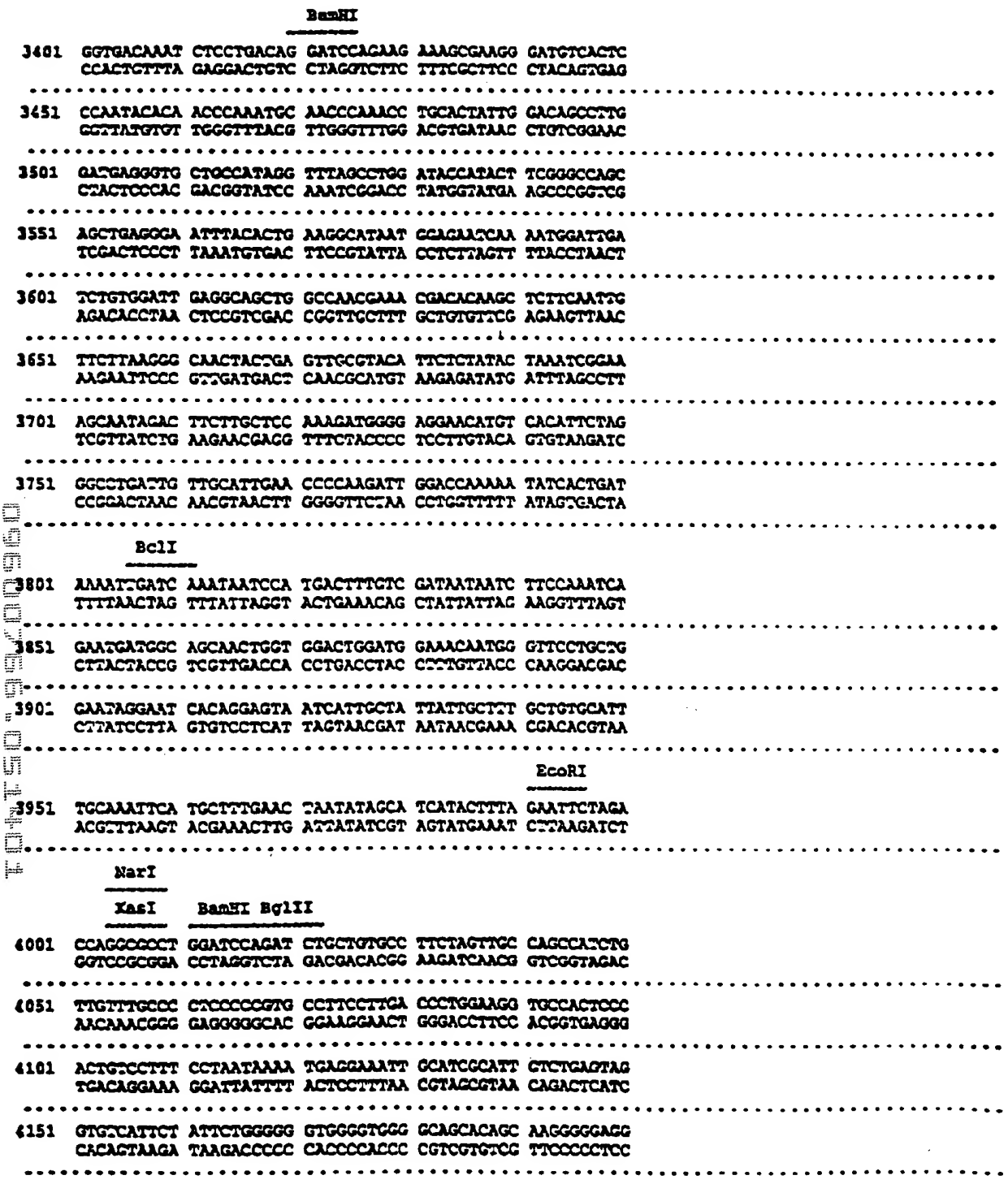
3151 CTGCCAAGAC CACCAGCCAA CCAACCAACA GCACAGAATC GACGACACTA
GACGGTTCTG GTGGTCGGTT GGTTGGTTGT CGTGTCTTAG CTGCTGTGAT

3201 AACCCAACAT CAGAGCCCTC CAGTAGAGGC ACGGGACCAT CCAGCCCCAC
TTGGGTTGTA GTCTCGGGAG GTCATCTCCG TGCCCTGGTA GGTCGGGGTG

3251 GGTCCCCAAC ACCACAGAAA GCCACGCCGA ACTTGGCAAG ACAACCCCAA
CCAGGGGTTG TGGTGTCTTT CGGTGCGGCT TGAACCGTTC TGTTGGGGTT

3301 CCACACTCCC AGAACAGCAC ACTGCCGCCA GTGCCATTCC AAGAGCCGTG
GGTGTGAGGG TCTTGTCGTG TGACGGCGGT CACGGTAAGG TTCTCGGCAC

3351 CACCCCGACG AACTCAGTGG ACCTGGCTTC CTGACGAACA CAATACGGGG
GTGGGGCTGC TTGAGTCACC TGGACCGAAG GACTGCTTGT GTTATGCCCC


BamHI
3401 GGTGACAAAT CTCCTGACAG GATCCAGAAG AAAGCGAAGG GATGTCACTC
CCACTGTTTA GAGGACTGTC CTAGGTCTTC TTTCGCTTCC CTACAGTGAG
3451 CCAATACACA ACCCAAATGC AACCCAAACC TGCACTATTG GACAGCCTTG
GGTTATGTGT TGGGTTTACG TTGGGTTTGG ACGTGATAAC CTGTCGGAAC
3501 GATGAGGGTG CTGCCATAGG TTTAGCCTGG ATACCATACT TCGGGCCAGC
CTACTCCCAC GACGGTATCC AAATCGGACC TATGGTATGA AGCCCGGTCG
3551 AGCTGAGGGA ATTTACACTG AAGGCATAAT GGAGAACTAA AATGGATTGA
TCGACTCCCT TAAATGTGAC TTCCGTATTA CCTCTTAGTT TTACCTAACT
3601 TCTGTGGATT GAGGCAGCTG GCCAACGAAA CGACACAAGC TCTTCAATTG
AGACACCTAA CTCCGTCGAC CGGTTGCTTT GCTGTGTTCG AGAAGTTAAC
3651 TTCTTAAGGG CAACTACTGA GTTGCGTACA TTCTCTATAC TAAATCGGAA
AAGAATTCCC GTTGATGACT CAACGCATGT AAGAGATATG ATTTAGCCTT
3701 AGCAATAGAC TTCTTGCTCC AAAGATGGGG AGGAACATGT CACATTCTAG
TCGTTATCTG AAGAACGAGG TTTCTACCCC TCCTTGTACA GTGTAAGATC
3751 GGCCTGATTG TTGCATTGAA CCCCAAGATT GGACCAAAAA TATCACTGAT
CCGGACTAAC AACGTAACTT GGGGTTCTAA CCTGGTTTTT ATAGTGACTA
BclI
3801 AAAATTGATC AAATAATCCA TGACTTTGTC GATAATAATC TTCCAAATCA
TTTTAACTAG TTTATTAGGT ACTGAAACAG CTATTATTAG AAGGTTTAGT
3851 GAATGATGGC AGCAACTGGT GGACTGGATG GAAACAATGG GTTCCTGCTG
CTTACTACCG TCGTTGACCA CCTGACCTAC CTTTGTTACC CAAGGACGAC
3901 GAATAGGAAT CACAGGAGTA ATCATTGCTA TTATTGCTTT GCTGTGCATT
CTTATCCTTA GTGTCCTCAT TAGTAACGAT AATAACGAAA CGACACGTAA
EcoRI
3951 TGCAAATTCA TGCTTTGAAC TAATATAGCA TCATACTTTA GAATTCTAGA
ACGTTTAAGT ACGAAACTTG ATTATATCGT AGTATGAAAT CTTAAGATCT
NarI
XasI BamHI BglII
4001 CCAGGCGCCT GGATCCAGAT CTGCTGTGCC TTCTAGTTGC CAGCCATCTG
GGTCCGCGGA CCTAGGTCTA GACGACACGG AAGATCAACG GTCGGTAGAC
4051 TTGTTTGCCC CTCCCCCGTG CCTTCCTTGA CCCTGGAAGG TGCCACTCCC
AACAAACGGG GAGGGGGCAC GGAAGGAACT GGGACCTTCC ACGGTGAGGG
4101 ACTGTCCTTT CCTAATAAAA TGAGGAAATT GCATCGCATT GTCTGAGTAG
TGACAGGAAA GGATTATTTT ACTCCTTTAA CGTAGCGTAA CAGACTCATC
4151 GTGTCATTCT ATTCTGGGGG GTGGGGTGGG GCAGCACAGC AAGGGGGAGG
CACAGTAAGA TAAGACCCCC CACCCCACCC CGTCGTGTCG TTCCCCCTCC

```
                          SphI                    KpnI
4201  ATTGGGAAGA CAATAGCAGG CATGCTGGGG ATGCGGTGGG CTCTATGGGT
      TAACCCTTCT GTTATCGTCC GTACGACCCC TACGCCACCC GAGATACCCA
........................................................................
      KpnI
4251  ACCCAGGTGC TGAAGAATTG ACCCGGTTCC TCCTGGGCCA GAAAGAAGCA
      TGGGTCCACG ACTTCTTAAC TGGGCCAAGG AGGACCCGGT CTTTCTTCGT
........................................................................
4301  GGCACATCCC CTTCTCTGTG ACACACCCTG TCCACGCCCC TGGTTCTTAG
      CCGTGTAGGG GAAGAGACAC TGTGTGGGAC AGGTGCGGGG ACCAAGAATC
........................................................................
4351  TTCCAGCCCC ACTCATAGGA CACTCATAGC TCAGGAGGGC TCCGCCTTCA
      AAGGTCGGGG TGAGTATCCT GTGAGTATCG AGTCCTCCCG AGGCGGAAGT
........................................................................
4401  ATCCCACCCC CTAAAGTACT TGGAGCGGTC TCTCCCTCCC TCATCAGCCC
      TAGGGTGGGG GATTTCATGA ACCTCGCCAG AGAGGGAGGG AGTAGTCGGG
........................................................................
4451  ACCAAACCAA ACCTAGCCTC CAAGAGTGGG AAGAAATTAA AGCAAGATAG
      TGGTTTGGTT TGGATCGGAG GTTCTCACCC TTCTTTAATT TCGTTCTATC
........................................................................
4501  GCTATTAAGT GCAGAGGGAG AGAAAATGCC TCCAACATGT GAGGAAGTAA
      CGATAATTCA CGTCTCCCTC TCTTTTACGG AGGTTGTACA CTCCTTCATT
........................................................................
                      XmnI
4551  TGAGAGAAAT CATAGAATTT CTTCCGCTTC CTCGCTCACT GACTCGCTGC
      ACTCTCTTTA GTATCTTAAA GAAGGCGAAG GAGCGAGTGA CTGAGCGACG
........................................................................
4601  GCTCGGTCGT TCGGCTGCGG CGAGCGGTAT CAGCTCACTC AAAGGCGGTA
      CGAGCCAGCA AGCCGACGCC GCTCGCCATA GTCGAGTGAG TTTCCGCCAT
........................................................................
4651  ATACGGTTAT CCACAGAATC AGGGGATAAC GCAGGAAAGA ACATGTGAGC
      TATGCCAATA GGTGTCTTAG TCCCCTATTG CGTCCTTTCT TGTACACTCG
........................................................................
4701  AAAAGGCCAG CAAAAGGCCA GGAACCGTAA AAAGGCCGCG TTGCTGGCGT
      TTTTCCGGTC GTTTTCCGGT CCTTGGCATT TTTCCGGCGC AACGACCGCA
........................................................................
4751  TTTTCCATAG GCTCCGCCCC CCTGACGAGC ATCACAAAAA TCGACGCTCA
      AAAAGGTATC CGAGGCGGGG GGACTGCTCG TAGTGTTTTT AGCTGCGAGT
........................................................................
4801  AGTCAGAGGT GGCGAAACCC GACAGGACTA TAAAGATACC AGGCGTTTCC
      TCAGTCTCCA CCGCTTTGGG CTGTCCTGAT ATTTCTATGG TCCGCAAAGG
........................................................................
4851  CCCTGGAAGC TCCCTCGTGC GCTCTCCTGT TCCGACCCTG CCGCTTACCG
      GGGACCTTCG AGGGAGCACG CGAGAGGACA AGGCTGGGAC GGCGAATGGC
........................................................................
4901  GATACCTGTC CGCCTTTCTC CCTTCGGGAA GCGTGGCGCT TTCTCAATGC
      CTATGGACAG GCGGAAAGAG GGAAGCCCTT CGCACCGCGA AAGAGTTACG
........................................................................
4951  TCACGCTGTA GGTATCTCAG TTCGGTGTAG GTCGTTCGCT CCAAGCTGGG
      AGTGCGACAT CCATAGAGTC AAGCCACATC CAGCAAGCGA GGTTCGACCC
........................................................................
5001  CTGTGTGCAC GAACCCCCCG TTCAGCCCGA CCGCTGCGCC TTATCCGGTA
      GACACACGTG CTTGGGGGGC AAGTCGGGCT GGCGACGCGG AATAGGCCAT
........................................................................
```

```
5051  ACTATCGTCT TGAGTCCAAC CCGGTAAGAC ACGACTTATC GCCACTGGCA
      TGATAGCAGA ACTCAGGTTG GGCCATTCTG TGCTGAATAG CGGTGACCGT
 .....
5101  GCAGCCACTG GTAACAGGAT TAGCAGAGCG AGGTATGTAG GCGGTGCTAC
      CGTCGGTGAC CATTGTCCTA ATCGTCTCGC TCCATACATC CGCCACGATG
 .....
5151  AGAGTTCTTG AAGTGGTGGC CTAACTACGG CTACACTAGA AGGACAGTAT
      TCTCAAGAAC TTCACCACCG GATTGATGCC GATGTGATCT TCCTGTCATA
 .....
5201  TTGGTATCTG CGCTCTGCTG AAGCCAGTTA CCTTCGGAAA AAGAGTTGGT
      AACCATAGAC GCGAGACGAC TTCGGTCAAT GGAAGCCTTT TTCTCAACCA
 .....
5251  AGCTCTTGAT CCGGCAAACA AACCACCGCT GGTAGCGGTG GTTTTTTTGT
      TCGAGAACTA GGCCGTTTGT TTGGTGGCGA CCATCGCCAC CAAAAAACA
 .....
5301  TTGCAAGCAG CAGATTACGC GCAGAAAAAA AGGATCTCAA GAAGATCCTT
      AACGTTCGTC GTCTAATGCG CGTCTTTTTT TCCTAGAGTT CTTCTAGGAA
 .....
5351  TGATCTTTTC TACGGGGTCT GACGCTCAGT GGAACGAAAA CTCACGTTAA
      ACTAGAAAAG ATGCCCCAGA CTGCGAGTCA CCTTGCTTTT GAGTGCAATT
 .....
5401  GGGATTTTGG TCATGAGATT ATCAAAAAGG ATCTTCACCT AGATCCTTTT
      CCCTAAAACC AGTACTCTAA TAGTTTTTCC TAGAAGTGGA TCTAGGAAAA
 .....
5451  AAATTAAAAA TGAAGTTTTA AATCAATCTA AAGTATATAT GAGTAAACTT
      TTTAATTTTT ACTTCAAAAT TTAGTTAGAT TTCATATATA CTCATTTGAA
 .....
5501  GGTCTGACAG TTACCAATGC TTAATCAGTG AGGCACCTAT CTCAGCGATC
      CCAGACTGTC AATGGTTACG AATTAGTCAC TCCGTGGATA GAGTCGCTAG
 .....
5551  TGTCTATTTC GTTCATCCAT AGTTGCCTGA CTCCGGGGGG GGGGGCGCT
      ACAGATAAAG CAAGTAGGTA TCAACGGACT GAGGCCCCCC CCCCCCGCGA
 .....
5601  GAGGTCTGCC TCGTGAAGAA GGTGTTGCTG ACTCATACCA GGCCTGAATC
      CTCCAGACGG AGCACTTCTT CCACAACGAC TGAGTATGGT CCGGACTTAG
 .....
5651  GCCCCATCAT CCAGCCAGAA AGTGAGGGAG CCACGGTTGA TGAGAGCTTT
      CGGGGTAGTA GGTCGGTCTT TCACTCCCTC GGTGCCAACT ACTCTCGAAA
 .....
5701  GTTGTAGGTG GACCAGTTGG TGATTTTGAA CTTTTGCTTT GCCACGGAAC
      CAACATCCAC CTGGTCAACC ACTAAAACTT GAAAACGAAA CGGTGCCTTG
 .....
5751  GGTCTGCGTT GTCGGGAAGA TGCGTGATCT GATCCTTCAA CTCAGCAAAA
      CCAGACGCAA CAGCCCTTCT ACGCACTAGA CTAGGAAGTT GAGTCGTTTT
 .....
5801  GTTCGATTTA TTCAACAAAG CCGCCGTCCC GTCAAGTCAG CGTAATGCTC
      CAACCTAAAT AAGTTGTTTC GGCGGCAGGG CAGTTCAGTC GCATTACGAG
 .....
5851  TGCCAGTGTT ACAACCAATT AACCAATTCT GATTAGAAAA ACTCATCGAG
      ACGGTCACAA TGTTGGTTAA TTGGTTAAGA CTAATCTTTT TGAGTAGCTC
 .....
5901  CATCAAATGA AACTGCAATT TATTCATATC AGGATTATCA ATACCATATT
      GTAGTTTACT TTGACGTTAA ATAAGTATAG TCCTAATAGT TATGGTATAA
 .....
5951  TTTGAAAAAG CCGTTTCTGT AATGAAGGAG AAAACTCACC GAGGCAGTTC
      AAACTTTTTC GGCAAAGACA TTACTTCCTC TTTTGAGTGG CTCCGTCAAG
 .....
```

```
6001  CATAGGATGG CAAGATCCTG GTATCGGTCT GCGATTCCGA CTCGTCCAAC
      GTATCCTACC GTTCTAGGAC CATAGCCAGA CGCTAAGGCT GAGCAGGTTG

6051  ATCAATACAA CCTATTAATT TCCCCTCGTC AAAAATAAGG TTATCAAGTG
      TAGTTATGTT GGATAATTAA AGGGGAGCAG TTTTTATTCC AATAGTTCAC

                                                 HindIII
6101  AGAAATCACC ATGAGTGACG ACTGAATCCG GTGAGAATGG CAAAAGCTTA
      TCTTTAGTGG TACTCACTGC TGACTTAGGC CACTCTTACC GTTTTCGAAT

6151  TGCATTTCTT TCCAGACTTG TTCAACAGGC CAGCCATTAC GCTCGTCATC
      ACGTAAAGAA AGGTCTGAAC AAGTTGTCCG GTCGGTAATG CGAGCAGTAG

6201  AAAATCACTC GCATCAACCA AACCGTTATT CATTCGTGAT TGCGCCTGAG
      TTTTAGTGAG CGTAGTTGGT TTGGCAATAA GTAAGCACTA ACGCGGACTC

                 PvuI
6251  CCAGACGAAA TACGCGATCG CTGTTAAAAG GACAATTACA AACAGGAATC
      GGTCTGCTTT ATGCGCTAGC GACAATTTTC CTGTTAATGT TTGTCCTTAG

6301  GAATGCAACC GGCGCAGGAA CACTGCCAGC GCATCAACAA TATTTTCACC
      CTTACGTTGG CCGCGTCCTT GTGACGGTCG CGTAGTTGTT ATAAAAGTGG

6351  TGAATCAGGA TATTCTTCTA ATACCTGGAA TGCTGTTTTC CCGGGGATCG
      ACTTAGTCCT ATAAGAAGAT TATGGACCTT ACGACAAAAG GGCCCCTAGC

6401  CAGTGGTGAG TAACCATGCA TCATCAGGAG TACGGATAAA ATGCTTGATG
      GTCACCACTC ATTGGTACGT AGTAGTCCTC ATGCCTATTT TACGAACTAC

6451  GTCGGAAGAG GCATAAATTC CGTCAGCCAG TTTAGTCTGA CCATCTCATC
      CAGCCTTCTC CGTATTTAAG GCAGTCGGTC AAATCAGACT GGTAGAGTAG

6501  TGTAACATCA TTGGCAACGC TACCTTTGCC ATGTTTCAGA AACAACTCTG
      ACATTGTAGT AACCGTTGCG ATGGAAACGG TACAAAGTCT TTGTTGAGAC

                             ClaI
6551  GCGCATCGGG CTTCCCATAC AATCGATAGA TTGTCGCACC TGATTGCCCG
      CGCGTAGCCC GAAGGGTATG TTAGCTATCT AACAGCGTGG ACTAACGGGC

6601  ACATTATCGC GAGCCCATTT ATACCCATAT AAATCAGCAT CCATGTTGGA
      TGTAATAGCG CTCGGGTAAA TATGGGTATA TTTAGTCGTA GGTACAACCT

                 XhoI
6651  ATTTAATCGC GGCCTCGAGC AAGACGTTTC CCGTTGAATA TGGCTCATAA
      TAAATTAGCG CCGGAGCTCG TTCTGCAAAG GGCAACTTAT ACCGAGTATT

6701  CACCCCTTGT ATTACTGTTT ATGTAAGCAG ACAGTTTTAT TGTTCATGAT
      GTGGGGAACA TAATGACAAA TACATTCGTC TGTCAAAATA ACAAGTACTA

                                                  DraIII
6751  GATATATTTT TATCTTGTGC AATGTAACAT CAGAGATTTT GAGACACAAC
      CTATATAAAA ATAGAACACG TTACATTGTA GTCTCTAAAA CTCTGTGTTG
```

DraIII

```
6801  GTGGCTTTCC CCCCCCCCCC ATTATTGAAG CATTTATCAG GGTTATTGTC
      CACCGAAAGG GGGGGGGGGG TAATAACTTC GTAAATAGTC CCAATAACAG
.....
6851  TCATGAGCGG ATACATATTT GAATGTATTT AGAAAAATAA ACAAATAGGG
      AGTACTCGCC TATGTATAAA CTTACATAAA TCTTTTTATT TGTTTATCCC
.....
6901  GTTCCGCGCA CATTTCCCCG AAAAGTGCCA CCTGACGTCT AAGAAACCAT
      CAAGGCGCGT GTAAAGGGGC TTTTCACGGT GGACTGCAGA TTCTTTGGTA
.....
6951  TATTATCATG ACATTAACCT ATAAAAATAG GCGTATCACG AGGCCCTTTC
      ATAATAGTAC TGTAATTGGA TATTTTTATC CGCATAGTGC TCCGGGAAAG
.....
7001  GTC
      CAG
.....
```

pVR 1012-GP(S)

General Descripti n

DNA pVR 1012-GP(S)
Local object
Created: 09/14/98 03:58PM
Last Modification Date: ? (no data)
length: 7073 bp
storage type: Basic
form: Circular
Comments

Restriction Map

BalI: 1 site TGGCCA / ACCGGT

BclI: 1 site TGATCA / ACTAGT

ClaI: 1 site ATCGAT / TAGCTA

DraIII: 1 site CACNNNGTG / GTGNNNCAC

HindIII: 1 site AAGCTT / TTCGAA

KasI: 1 site GGCGCC / CCGCGG

KpnI: 1 site GGTACC / CCATGG

NarI: 1 site GGCGCC / CCGCGG

PmlI: 1 site CACGTG / GTGCAC

PvuI: 1 site CGATCG / GCTAGC

SacII: 1 site CCGCGG / GGCGCC

SalI: 1 site GTCGAC / CAGCTG

XbaI: 1 site TCTAGA / AGATCT

XmnI: 1 site GAANNNNTTC / CTTNNNNAAG

NdeI: 2 sites CATATG / GTATAC

EcoRV: 3 sites GATATC / CTATAG

SphI: 3 sites GCATGC / CGTACG

NcoI: 4 sites CCATGG / GGTACC

BamHI: 6 sites GGATCC / CCTAGG

Functional Map

CDS (4 signals)

CMV IE 5' UT

Start: 886 End: 1129

CMV IE INT

Start: 1130 End: 1840

TbGH

Start: 4090 End: 4642

Kan r

Start: 6138 End: 6760 (Complementary)

Misc_feature (2 signals)

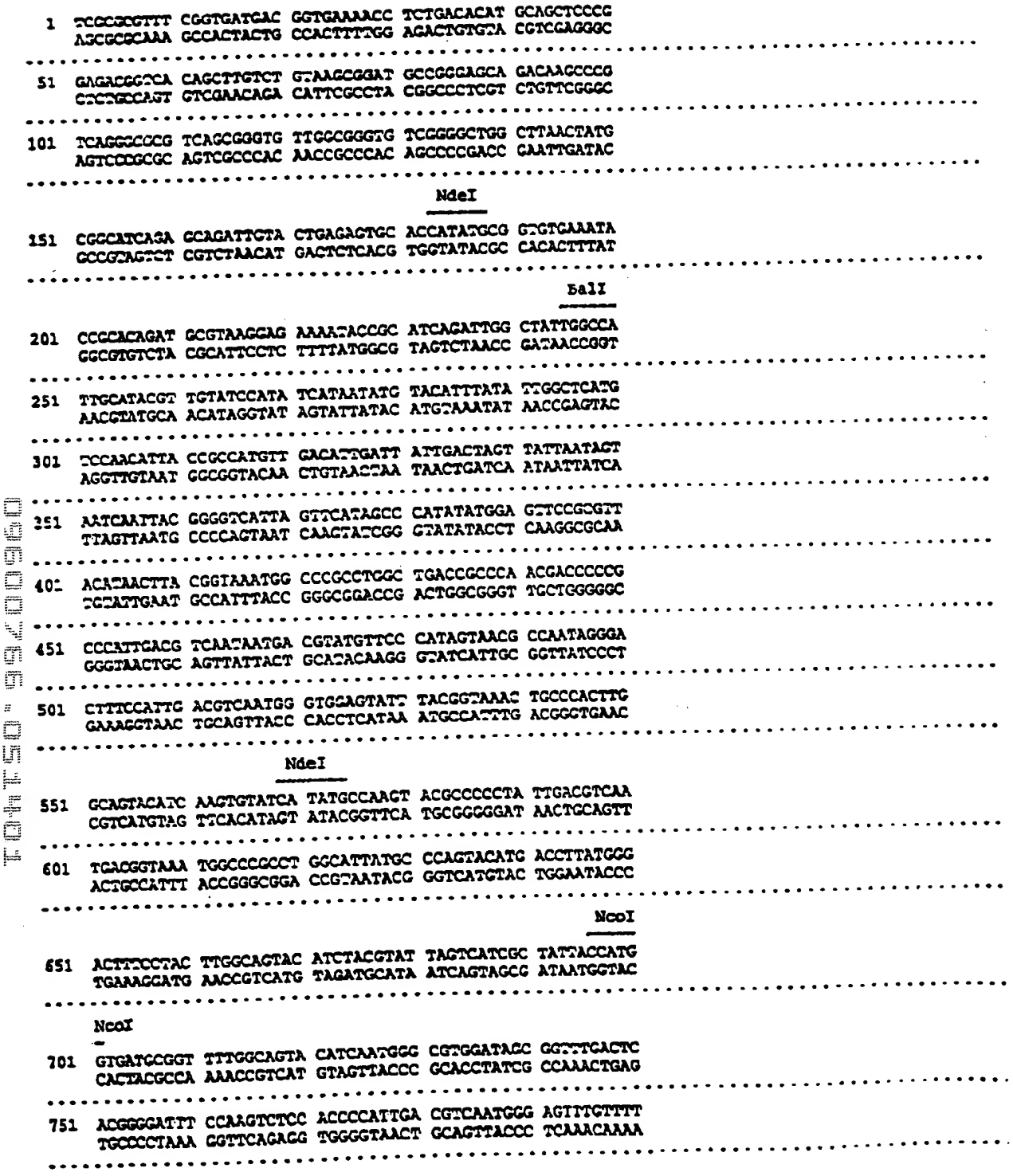
CMV enhancer

Start: 248 End: 885

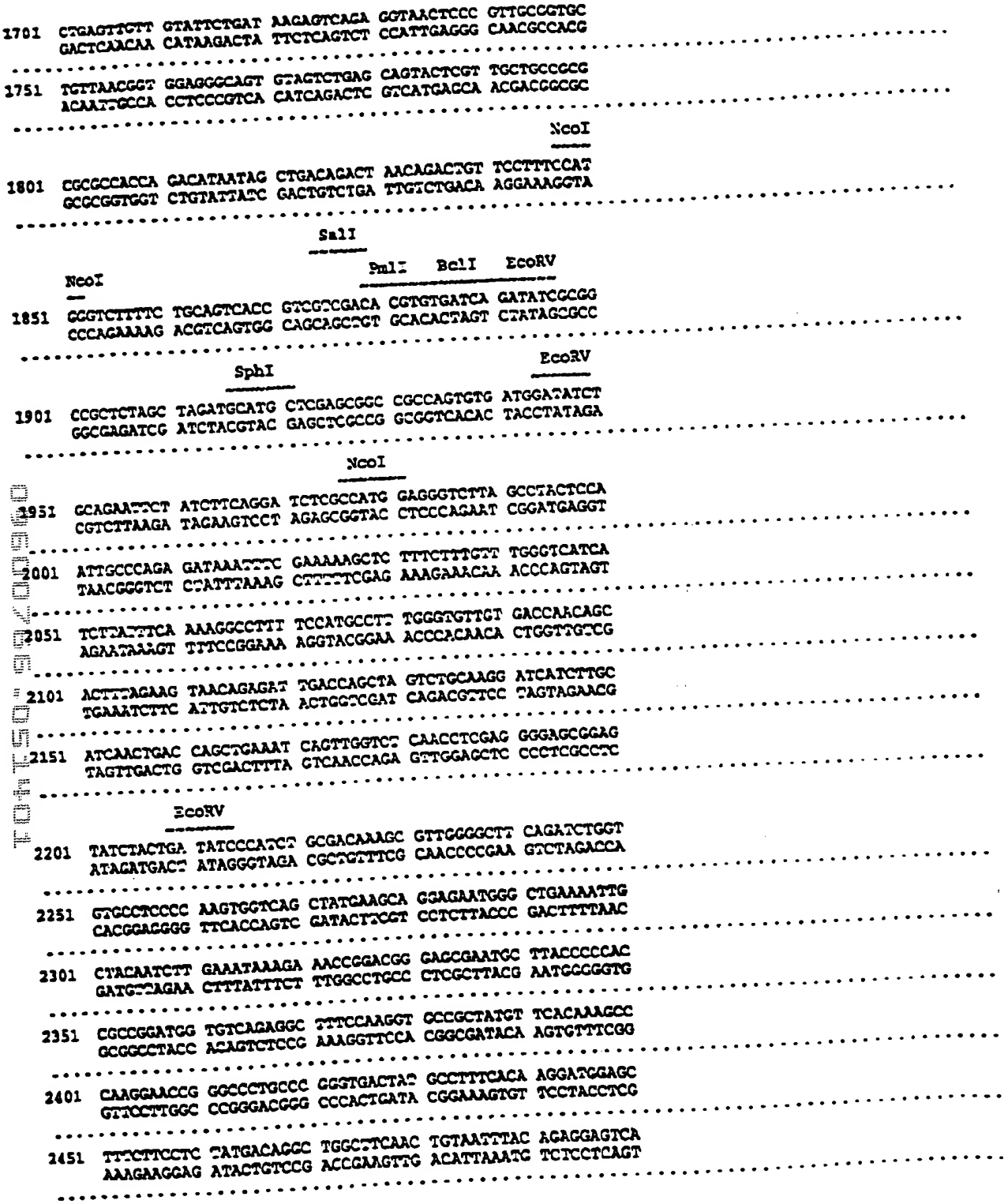
GP(S)

Start: 1870 End: 4089

Annotations


NdeI
BalI
NdeI
NcoI
NcoI

```
 801  GGCACCAAAA TCAACGGGAC TTTCCAAAAT GTCGTAACAA CTCCGCCCCA
      CCGTGGTTTT AGTTGCCCTG AAAGGTTTTA CAGCATTGTT GAGGCGGGGT
 851  TTGACCCAAA TGGGCGGTAG GCGTGTACGG TGGGAGGTCT ATATAAGCAG
      AACTGCGTTT ACCCGCCATC CGCACATGCC ACCCTCCAGA TATATTCGTC
 901  AGCTCGTTTA GTGAACCGTC AGATCGCCTG GAGACGCCAT CCACGCTGTT
      TCGAGCAAAT CACTTGGCAG TCTAGCGGAC CTCTGCGGTA GGTGCGACAA
                                          SacII
 951  TTGACCTCCA TAGAAGACAC CGGGACCGAT CCAGCCTCCG CGGCCGGGAA
      AACTGGAGGT ATCTTCTGTG GCCCTGGCTA GGTCGGAGGC GCCGGCCCTT
1001  CGGTGCATTG GAACGCGGAT TCCCCGTGCC AAGAGTGACG TAAGTACCGC
      GCCACGTAAC CTTGCGCCTA AGGGGCACGG TTCTCACTGC ATTCATGGCG
                                          SphI
1051  CTATAGACTC TATAGGCACA CCCCTTTGGC TCTTATGCAT GCTATACTGT
      GATATCTGAG ATATCCGTGT GGGGAAACCG AGAATACGTA CGATATGACA
1101  TTTTGGCTTG GGGCCTATAC ACCCCCGCTT CCTTATGCTA TAGGTGATGG
      AAAACCGAAC CCCGGATATG TGGGGGCGAA GGAATACGAT ATCCACTACC
1151  TATAGCTTAG CCTATAGGTG TGGGTTATTG ACCATTATTG ACCACTCCCC
      ATATCGAATC GGATATCCAC ACCCAATAAC TGGTAATAAC TGGTGAGGGG
1201  TATTCGTGAC GATACTTTCC ATTACTAATC CATAACATGG CTCTTTGCCA
      ATAACCACTG CTATGAAAGG TAATGATTAG GTATTGTACC GAGAAACGGT
1251  CAACTATCCC TATTGGCTAT ATGCCAATAC TCTGTCCTTC AGAGACTGAC
      GTTGATAGAG ATAACCGATA TACGGTTATG AGACAGGAAG TCTCTGACTG
1301  ACGGACTCTG TATTTTTACA GGATGGGGTC CCATTTATTA TTTACAAATT
      TGCCTGAGAC ATAAAAATGT CCTACCCCAG GGTAAATAAT AAATGTTTAA
1351  CACATATACA ACAACGCCGT CCCCCGTGCC CGCAGTTTTT ATTAAACATA
      GTGTATATGT TGTTGCGGCA GGGGGCACGG GCGTCAAAAA TAATTTGTAT
1401  GCGTGGGATC TCCACGCGAA TCTCGGGTAC GTGTTCCGGA CATGGGCTCT
      CGCACCCTAG AGGTGCGCTT AGAGCCCATG CACAAGGCCT GTACCCGAGA
1451  TCTCCGGTAG CGGCGGAGCT TCCACATCCG AGCCCTGGTC CCATGCCTCC
      AGAGGCCATC GCCGCCTCGA AGGTGTAGGC TCGGGACCAG GGTACGGAGG
1501  ACCGGCTCAT GGTCGCTCGG CAGCTCCTTG CTCCTAACAG TGGAGGCCAG
      TCGCCGAGTA CCAGCGAGCC GTCGAGGAAC GAGGATTGTC ACCTCCGGTC
1551  ACTTAGGCAC AGCACAATGC CCACCACCAC CAGTGTGCCG CACAAGGCCG
      TGAATCCGTG TCGTGTTACG GGTGGTGGTG GTCACACGGC GTGTTCCGGC
1601  TGGCGGTAGG GTATGTGTCT GAAAATGAGC GTGGAGATTG GGCTCGCACG
      ACCGCCATCC CATACACAGA CTTTTACTCG CACCTCTAAC CCGAGCGTGC
1651  GCTGACGCAG ATGGAAGACT TAAGGCAGCG GCAGAAGAAG ATGCAGGCAG
      CGACTGCGTC TACCTTCTGA ATTCCGTCGC CGTCTTCTTC TACGTCCGTC
```


1701 CTGAGTTGTT GTATTCTGAT AAGAGTCAGA GGTAACTCCC GTTGCGGTGC
GACTCAACAA CATAAGACTA TTCTCAGTCT CCATTGAGGG CAACGCCACG
1751 TGTTAACGGT GGAGGGCAGT GTAGTCTGAG CAGTACTCGT TGCTGCCGCG
ACAATTGCCA CCTCCCGTCA CATCAGACTC GTCATGAGCA ACGACGGCGC
NcoI
1801 CGCGCCACCA GACATAATAG CTGACAGACT AACAGACTGT TCCTTTCCAT
GCGCGGTGGT CTGTATTATC GACTGTCTGA TTGTCTGACA AGGAAAGGTA
SalI
PmlI BclI EcoRV
NcoI
1851 GGGTCTTTTC TGCAGTCACC GTCGTCGACA CGTGTGATCA GATATCGCGG
CCCAGAAAAG ACGTCAGTGG CAGCAGCTGT GCACACTAGT CTATAGCGCC
SphI
EcoRV
1901 CCGCTCTAGC TAGATGCATG CTCGAGCGGC CGCCAGTGTG ATGGATATCT
GGCGAGATCG ATCTACGTAC GAGCTCGCCG GCGGTCACAC TACCTATAGA
NcoI
1951 GCAGAATTCT ATCTTCAGGA TCTCGCCATG GAGGGTCTTA GCCTACTCCA
CGTCTTAAGA TAGAAGTCCT AGAGCGGTAC CTCCCAGAAT CGGATGAGGT
2001 ATTGCCCAGA GATAAATTTC GAAAAAGCTC TTTCTTTGTT TGGGTCATCA
TAACGGGTCT CTATTTAAAG CTTTTTCGAG AAAGAAACAA ACCCAGTAGT
2051 TCTTATTTCA AAAGGCCTTT TCCATGCCTT TGGGTGTTGT GACCAACAGC
AGAATAAAGT TTTCCGGAAA AGGTACGGAA ACCCACAACA CTGGTTGTCG
2101 ACTTTAGAAG TAACAGAGAT TGACCAGCTA GTCTGCAAGG ATCATCTTGC
TGAAATCTTC ATTGTCTCTA ACTGGTCGAT CAGACGTTCC TAGTAGAACG
2151 ATCAACTGAC CAGCTGAAAT CAGTTGGTCT CAACCTCGAG GGGAGCGGAG
TAGTTGACTG GTCGACTTTA GTCAACCAGA GTTGGAGCTC CCCTCGCCTC
EcoRV
2201 TATCTACTGA TATCCCATCT GCGACAAAGC GTTGGGGCTT CAGATCTGGT
ATAGATGACT ATAGGGTAGA CGCTGTTTCG CAACCCCGAA GTCTAGACCA
2251 GTGCCTCCCC AAGTGGTCAG CTATGAAGCA GGAGAATGGG CTGAAAATTG
CACGGAGGGG TTCACCAGTC GATACTTCGT CCTCTTACCC GACTTTTAAC
2301 CTACAATCTT GAAATAAAGA AACCGGACGG GACCGAATGC TTACCCCCAC
GATGTTAGAA CTTTATTTCT TTGGCCTGCC CTCGCTTACG AATGCGGGTG
2351 CGCCGGATGG TGTCAGAGGC TTTCCAAGGT GCCGCTATGT TCACAAAGCC
GCGGCCTACC ACAGTCTCCG AAAGGTTCCA CGGCGATACA AGTGTTTCGG
2401 CAAGGAACCG GGCCCTGCCC GGGTGACTAC GCCTTTCACA AGGATGGAGC
GTTCCTTGGC CCGGGACGGG CCCACTGATA CGGAAAGTGT TCCTACCTCG
2451 TTTCTTCCTC TATGACAGGC TGGCCTCAAC TGTAATTTAC AGAGGAGTCA
AAAGAAGGAG ATACTGTCCG ACCGAAGTTG ACATTAAATG TCTCCTCAGT

```
2501  ATTTCCCTGA GGGGGTAATC GCATTCTTGA TATTGGCTAA ACCAAAGGAA
      TAAAACGACT CCCCCATTAG CGTAAGAACT ATAACCGATT TGGTTTCCTT
.....
2551  ACGTTCCTTC AATCACCCCC CATTCGAGAG GCAGCAAACT ACACTGAAAA
      TGCAAGGAAG TTAGTGGGGG GTAAGCTCTC CGTCGTTTGA TGTGACTTTT
.....
2601  TACATCAAGT TACTATGCCA CATCCTACTT GGAGTACGAA ATCGAAAATT
      ATGTAGTTCA ATGATACGGT GTAGGATGAA CCTCATGCTT TAGCTTTTAA
.....
2651  TTGGTGCTCA ACACTCCACG ACCCTTTTCA AAATTAACAA TAATACTTTT
      AACCACGAGT TGTGAGGTGC TGGGAAAAGT TTTAATTGTT ATTATGAAAA
.....
2701  GTTCTTCTGG ACAGGCCCCA CACGCCTCAG TTCCTTTTCC AGCTGAATGA
      CAAGAAGACC TGTCCGGGGT GTGCGGAGTC AAGGAAAAGG TCGACTTACT
.....
2751  TACCATTCAA CTTCACCAAC AGTTGAGCAA CACAACTGGG AAACTAATTT
      ATGGTAAGTT GAAGTGGTTG TCAACTCGTT GTGTTGACCC TTTGATTAAA
.....
2801  GGACACTAGA TGCTAATATC AATGCTGATA TTCGTGAATG GGCTTTTTGG
      CCTGTGATCT ACGATTATAG TTACGACTAT AACCACTTAC CCGAAAAACC
.....
2851  GAAAATAAAA AAATCTCTCC GAACAACTAC GTGGAGAAGA GCTGTCTTTC
      CTTTTATTTT TTTAGAGAGG CTTGTTGATG CACCTCTTCT CGACAGAAAG
.....
2901  GAAACTTTAT CGCTCAACGA GACAGAAGAC GATGATGCGA CATCGTCGAG
      CTTTGAAATA GCGAGTTGCT CTGTCTTCTG CTACTACGCT GTAGCAGCTC
.....
2951  AACTACAAAG GGAAGAATCT CCGACCGGGC CACCAGGAAG TATTCGGACC
      TTGATGTTTC CCTTCTTAGA GGCTGGCCCG GTGGTCCTTC ATAAGCCTGG
.....
3001  TGGTTCCAAA GGATTCCCCT GGGATGGTTT CATTGCACGT ACCAGAAGGG
      ACCAAGGTTT CCTAAGGGGA CCCTACCAAA GTAACGTGCA TGGTCTTCCC
.....
3051  GAAACAACAT TGCCGTCTCA GAATTCGACA GAAGGTCGAA GAGTAGATGT
      CTTTGTTGTA ACGGCAGAGT CTTAAGCTGT CTTCCAGCTT CTCATCTACA
.....
3101  GAATACTCAG GAAACTATCA CAGAGACAAC TGCAACAATC ATAGGCACTA
      CTTATGAGTC CTTTGATAGT GTCTCTGTTG ACGTTGTTAG TATCCGTGAT
.....
3151  ACGGTAACAA CATGCAGATC TCCACCATCG GGACAGGACT GAGCTCCAGC
      TGCCATTGTT GTACGTCTAG AGGTGGTAGC CCTGTCCTGA CTCGAGGTCG
.....
                          NcoI
                          ------
3201  CAAATCCTGA GTTCCTCACC GACCATGGCA CCAAGCCCTG AGACTCAGAC
      GTTTAGGACT CAAGGAGTGG CTGGTACCGT GGTTCGGGAC TCTGAGTCTG
.....
3251  CTCCACAACC TACACACCAA AACTACCAGT GATGACCACC GAGGAACCAA
      GAGGTGTTGG ATGTGTGGTT TTGATGGTCA CTACTGGTGG CTCCTTGGTT
.....
3301  CAACACCACC GAGAAACTCT CCTGGCTCAA CAACAGAAGC ACCCACTCTC
      GTTGTGGTGG CTCTTTGAGA GGACCGAGTT GTTGTCTTCG TGGGTGAGAG
.....
3351  ACCACCCCAG AGAATATAAC AACAGCGGTT AAAACTGTTT GGGCACAAGA
      TGGTGGGGTC TCTTATATTG TTGTCGCCAA TTTTGACAAA CCCGTGTTCT
.....
```

```
3401  GTCCACAAGC AACGGTCTAA TAACTTCAAC AGTAACAGGT ATTCTTGGGA
      CAGGTGTTCG TTGCCAGATT ATTGAAGTTG TCATTGTCCA TAAGAACCCT
.....
3451  GCCTTGGACT TCGAAAACGC AGCAGAAGAC AAGTTAACAC CAGGGCCACG
      CGGAACCTGA AGCTTTTGCG TCGTCTTCTG TTCAATTGTG GTCCCGGTGC
.....
3501  GGTAAATGCA ATCCCAACTT ACACTACTGG ACTGCACAAG AACAACATAA
      CCATTTACGT TAGGGTTGAA TGTGATGACC TGACGTGTTC TTGTTGTATT
.....
                      BamHI
                      ------
3551  TGCTGCTGGG ATTGCCTGGA TCCCGTACTT TGCACCGGGT GCAGAAGGCA
      ACGACGACCC TAACGGACCT AGGGCATGAA ACCTGGCCCA CGTCTTCCGT
.....
3601  TATACACTGA AGGCCTTATG CACAACCAAA ATGCCTTAGT CTGTGGACTC
      ATATGTGACT TCCGGAATAC GTGTTGGTTT TACGGAATCA GACACCTGAG
.....
3651  AGACAACTTG CAAATGAAAC AACTCAAGCT CTGCAGCTTT TCTTAAGGGC
      TCTGTTGAAC GTTTACTTTG TTGAGTTCGA GACGTCGAAA AGAATTCCCG
.....
3701  CACGACGGAG CTGCCGACAT ATACCATACT CAATAGGAAG GCCATAGATT
      GTGCTGCCTC GACGCCTGTA TATGGTATGA GTTATCCTTC CGGTATCTAA
.....
                                      BamHI
                                      ------
3751  TCCTTCTGCG ACGATGGGCC GGGACATGTA GGATCCTGGG ACCAGATTGT
      AGGAAGACGC TGCTACCCCG CCCTGTACAT CCTAGGACCC TGGTCTAACA
.....
3801  TGCATTGAGC CACATCATTG GACCAAAAAC ATCACTGATA AAATCAACCA
      ACGTAACTCG GTGTACTAAC CTGGTTTTTG TAGTGACTAT TTTAGTTGGT
.....
3851  AATCATCCAT GATTTCATCG ACAACCCTTT ACCCAATCAG GATAATGATG
      TTAGTAGGTA CTAAAGTAGC TGTTGGGAAA TGGGTTAGTC CTATTACTAC
.....
                                   BamHI
                                   ------
3901  ATAATTGGTG GACGGGCTGG AGACAGTGGA TCCCTGCAGG AATAGGCATT
      TATTAACCAC CTGCCCGACC TCTGTCACCT AGGGACGTCC TTATCCGTAA
.....
3951  ACTGGAATTA TTATTGCAAT CATTGCTCTT CTTTGCGTCT GCAAGCTGCT
      TGACCTTAAT AATAACGTTA GTAACGAGAA GAAACGCAGA CGTTCGACGA
.....
                                                  BamHI
                                                  ------
4001  TTGTTGAATA TCAGAATTCC AGCACTGGCG GCCGTTACTA GTGGATCCGA
      AACAACTTAT AGTCTTAAGG TCGTGACCGC CGGCAATGAT CACCTAGGCT
.....
                                NarI
                                ----
          BamHI       XbaI      KasI    BamHI
          ------      ----      ----    ------
4051  GCTCGGATCC AAGCTCTAGA CCAGGCGCCT GGATCCAGAT CTGCTGTGCC
      CGAGCCTAGG TTCGAGATCT GGTCCGCGGA CCTAGGTCTA GACGACACGG
.....
4101  TTCTAGTTGC CAGCCATCTG TTGTTTGCCC CTCCCCCGTG CCTTCCTTGA
      AAGATCAACG GTCGGTAGAC AACAAACGGG GAGGGGGCAC GGAAGGAACT
.....
4151  CCCTGGAAGG TGCCACTCCC ACTGTCCTTT CCTAATAAAA TGAGGAAATT
      GGGACCTTCC ACGGTGAGGG TGACAGGAAA GGATTATTTT ACTCCTTTAA
.....
```

```
4201  GCATCGCATT CTCTGAGTAG GTGTCATTCT ATTCTGGGGG GTGGGGTGGG
      CGTAGCGTAA GAGACTCATC CACAGTAAGA TAAGACCCCC CACCCCACCC

                                                SphI
4251  GCAGCACAGC AAGGGGGAGG ATTGGGAAGA CAATAGCAGG CATGCTGGGG
      CGTCGTGTCG TTCCCCCTCC TAACCCTTCT GTTATCGTCC GTACGACCCC

                         KpnI
4301  ATGCGGTGGG CTCTATGGGT ACCCAGGTGC TGAAGAATTG ACCCGGTTCC
      TACGCCACCC GAGATACCCA TGGGTCCACG ACTTCTTAAC TGGGCCAAGG

4351  TCCTGGGCCA GAAAGAAGCA GGCACATCCC CTTCTCTGTG ACACACCCTG
      AGGACCCGGT CTTTCTTCGT CCGTGTAGGG GAAGAGACAC TGTGTGGGAC

4401  TCCACGCCCC TGGTTCTTAG TTCCAGCCCC ACTCATAGGA CACTCATAGC
      AGGTGCGGGG ACCAAGAATC AAGGTCGGGG TGAGTATCCT GTGAGTATCG

4451  TCAGGAGGGC TCCGCCTTCA ATCCCACCCG CTAAAGTACT TGGAGCGGTC
      AGTCCTCCCG AGGCGGAAGT TAGGGTGGGC GATTTCATGA ACCTCGCCAG

4501  TCTCCCTCCC TCATCAGCCC ACCAAACCAA ACCTAGCCTC CAAGAGTGGG
      AGAGGGAGGG AGTAGTCGGG TGGTTTGGTT TGGATCGGAG GTTCTCACCC

4551  AAGAAATTAA AGCAAGATAG GCTATTAAGT GCAGAGGGAG AGAAAATGCC
      TTCTTTAATT TCGTTCTATC CGATAATTCA CGTCTCCCTC TCTTTTACGG

                                            XmnI
4601  TCCAACATGT GAGGAAGTAA TGAGAGAAAT CATAGAATTT CTTCCGCTTC
      AGGTTGTACA CTCCTTCATT ACTCTCTTTA GTATCTTAAA GAAGGCGAAG

4651  CTCGCTCACT GACTCGCTGC GCTCGGTCGT TCGGCTGCGG CGAGCGGTAT
      GAGCGAGTGA CTGAGCGACG CGAGCCAGCA AGCCGACGCC GCTCGCCATA

4701  CAGCTCACTC AAAGGCGGTA ATACGGTTAT CCACAGAATC AGGGGATAAC
      GTCGAGTGAG TTTCCGCCAT TATGCCAATA GGTGTCTTAG TCCCCTATTG

4751  GCAGGAAAGA ACATGTGAGC AAAAGGCCAG CAAAAGGCCA GGAACCGTAA
      CGTCCTTTCT TGTACACTCG TTTTCCGGTC GTTTTCCGGT CCTTGGCATT

4801  AAAGGCCGCG TTGCTGGCGT TTTTCCATAG GCTCCGCCCC CCTGACGAGC
      TTTCCGGCGC AACGACCGCA AAAAGGTATC CGAGGCGGGG GGACTGCTCG

4851  ATCACAAAAA TCGACGCTCA AGTCAGAGGT GGCGAAACCC GACAGGACTA
      TAGTGTTTTT AGCTGCGAGT TCAGTCTCCA CCGCTTTGGG CTGTCCTGAT

4901  TAAAGATACC AGGCGTTTCC CCCTGGAAGC TCCCTCGTGC GCTCTCCTGT
      ATTTCTATGG TCCGCAAAGG GGGACCTTCG AGGGAGCACG CGAGAGGACA

4951  TCCGACCCTG CCGCTTACCG GATACCTGTC CGCCTTTCTC CCTTCGGGAA
      AGGCTGGGAC GGCGAATGGC CTATGGACAG GCGGAAAGAG GGAAGCCCTT

5001  GCGTGGCGCT TTCTCAATGC TCACGCTGTA GGTATCTCAG TTCGGTGTAG
      CGCACCGCGA AAGAGTTACG AGTGCGACAT CCATAGAGTC AAGCCACATC
```

```
5051  GTCGTTCGCT CCAAGCTGGG CTGTGTGCAC GAACCCCCCG TTCAGCCCGA
      CAGCAAGCGA GGTTCGACCC GACACACGTG CTTGGGGGGC AAGTCGGGCT
.....
5101  CCGCTGCGCC TTATCCGGTA ACTATCGTCT TGAGTCCAAC CCGGTAAGAC
      GGCGACGCGG AATAGGCCAT TGATAGCAGA ACTCAGGTTG GGCCATTCTG
.....
5151  ACGACTTATC GCCACTGGCA GCAGCCACTG GTAACAGGAT TAGCAGAGCG
      TGCTGAATAG CGGTGACCGT CGTCGGTGAC CATTGTCCTA ATCGTCTCGC
.....
5201  AGGTATGTAG GCGGTGCTAC AGAGTTCTTG AAGTGGTGGC CTAACTACGG
      TCCATACATC CGCCACGATG TCTCAAGAAC TTCACCACCG GATTGATGCC
.....
5251  CTACACTAGA AGGACAGTAT TTGGTATCTG CGCTCTGCTG AAGCCAGTTA
      GATGTGATCT TCCTGTCATA AACCATAGAC GCGAGACGAC TTCGGTCAAT
.....
5301  CCTTCGGAAA AAGAGTTGGT AGCTCTTGAT CCGGCAAACA AACCACCGCT
      GGAAGCCTTT TTCTCAACCA TCGAGAACTA GGCCGTTTGT TTGGTGGCGA
.....
5351  GGTAGCGGTG GTTTTTTTGT TTGCAAGCAG CAGATTACGC GCAGAAAAAA
      CCATCGCCAC CAAAAAAACA AACGTTCGTC GTCTAATGCG CGTCTTTTTT
.....
5401  AGGATCTCAA GAAGATCCTT TGATCTTTTC TACGGGGTCT GACGCTCAGT
      TCCTAGAGTT CTTCTAGGAA ACTAGAAAAG ATGCCCCAGA CTGCGAGTCA
.....
5451  GGAACGAAAA CTCACGTTAA GGGATTTTGG TCATGAGATT ATCAAAAAGG
      CCTTGCTTTT GAGTGCAATT CCCTAAAACC AGTACTCTAA TAGTTTTTCC
.....
5501  ATCTTCACCT AGATCCTTTT AAATTAAAAA TGAAGTTTTA AATCAATCTA
      TAGAAGTGGA TCTAGGAAAA TTTAATTTTT ACTTCAAAAT TTAGTCAGAT
.....
5551  AAGTATATAT GAGTAAACTT GGTCTGACAG TTACCAATGC TTAATCAGTG
      TTCATATATA CTCATTTGAA CCAGACTGTC AATGGTTACG AATTAGTCAC
.....
5601  AGGCACCTAT CTCAGCGATC TGTCTATTTC GTTCATCCAT AGTTGCCTGA
      TCCGTGGATA GAGTCGCTAG ACAGATAAAG CAAGTAGGTA TCAACGGACT
.....
5651  CTCCGGGGGG GGGGGGCGCT GAGGTCTGCC TCGTGAAGAA GGTGTTGCTG
      GAGGCCCCCC CCCCCCGCGA CTCCAGACGG AGCACTTCTT CCACAACGAC
.....
5701  ACTCATACCA GGCCTGAATC GCCCCATCAT CCAGCCAGAA AGTGAGGGAG
      TGAGTATGGT CCGGACTTAG CGGGGTAGTA GGTCGGTCTT TCACTCCCTC
.....
5751  CCACGGTTGA TGAGAGCTTT GTTGTAGGTG GACCAGTTGG TGATTTTGAA
      GGTGCCAACT ACTCTCGAAA CAACATCCAC CTGGTCAACC ACTAAAACTT
.....
5801  CTTTTGCTTT GCCACGGAAC GGTCTGCGTT GTCGGGAAGA TGCGTGATCT
      GAAAACGAAA CGGTGCCTTG CCAGACGCAA CAGCCCTTCT ACGCACTAGA
.....
5851  GATCCTTCAA CTCAGCAAAA GTTCGATTTA TTCAACAAAG CCGCCGTCCC
      CTAGGAAGTT GAGTCGTTTT CAAGCTAAAT AAGTTGTTTC GGCGGCAGGG
.....
5901  GTCAAGTCAG CGTAATGCTC TGCCAGTGTT ACAACCAATT AACCAATTCT
      CAGTTCAGTC GCATTACGAG ACGGTCACAA TGTTGGTTAA TTGGTTAAGA
.....
5951  GATTAGAAAA ACTCATCGAG CATCAAATGA AACTGCAATT TATTCATATC
      CTAATCTTTT TGAGTAGCTC GTAGTTTACT TTGACGTTAA ATAAGTATAG
.....
```

```
6001  AGGATTATCA ATACCATATT TTTGAAAAAG CCGTTTCTGT AATGAAGGAG
      TCCTAATAGT TATGGTATAA AAACTTTTTC GGCAAAGACA TTACTTCCTC

6051  AAAACTCACC GAGGCAGTTC CATACGATGG CAAGATCCTG GTATCGGTCT
      TTTTGAGTGG CTCCGTCAAG GTATCCTACC GTTCTAGGAC CATAGCCAGA

6101  GCGATTCCGA CTCGTCCAAC ATCAATACAA CCTATTAATT TCCCCTCGTC
      CGCTAAGGCT GAGCAGGTTG TAGTTATGTT GGATAATTAA AGGGGAGCAG

6151  AAAAATAAGG TTATCAAGTG AGAAATCACC ATGAGTGACG ACTGAATCCG
      TTTTTATTCC AATAGTTCAC TCTTTAGTGG TACTCACTGC TGACTTAGGC

                  HindIII
                  ═══════
6201  GTGAGAATGG CAAAAGCTTA TGCATTTCTT TCCAGACTTG TTCAACAGGC
      CACTCTTACC GTTTTCGAAT ACGTAAAGAA AGGTCTGAAC AAGTTGTCCG

6251  CAGCCATTAC GCTCGTCATC AAAATCACTC GCATCAACCA AACCGTTATT
      GTCGGTAATG CGAGCAGTAG TTTTAGTGAG CGTAGTTGGT TTGGCAATAA

                                        PvuI
                                        ════
6301  CATTCGTGAT TGCGCCTGAG CGAGACGAAA TACGCGATCG CTGTTAAAAG
      GTAAGCACTA ACGCGGACTC GCTCTGCTTT ATGCGCTAGC GACAATTTTC

6351  GACAATTACA AACAGGAATC GAATGCAACC GGCGCAGGAA CACTGCCAGC
      CTGTTAATGT TTGTCCTTAG CTTACGTTGG CCGCGTCCTT GTGACGGTCG

6401  GCATCAACAA TATTTTCACC TGAATCAGGA TATTCTTCTA ATACCTGGAA
      CGTAGTTGTT ATAAAAGTGG ACTTAGTCCT ATAAGAAGAT TATGGACCTT

6451  TGCTGTTTTC CCGGGGATCG CAGTGGTGAG TAACCATGCA TCATCAGGAG
      ACGACAAAAG GGCCCCTAGC GTCACCACTC ATTGGTACGT AGTAGTCCTC

6501  TACGGATAAA ATGCTTGATG GTCGGAAGAG GCATAAATTC CGTCAGCCAG
      ATGCCTATTT TACGAACTAC CAGCCTTCTC CGTATTTAAG GCAGTCGGTC

6551  TTTAGTCTGA CCATCTCATC TGTAACATCA TTGGCAACGC TACCTTTGCC
      AAATCAGACT GGTAGAGTAG ACATTGTAGT AACCGTTGCG ATGGAAACGG

                                                  ClaI
                                                  ════
6601  ATGTTTCAGA AACAACTCTG GCGCATCGGG CTTCCCATAC AATCGATAGA
      TACAAAGTCT TTGTTGAGAC CGCGTAGCCC GAAGGGTATG TTAGCTATCT

6651  TTGTCGCACC TGATTGCCCG ACATTATCGC GAGCCCATTT ATACCCATAT
      AACAGCGTGG ACTAACGGGC TGTAATAGCG CTCGGGTAAA TATGGGTATA

6701  AAATCAGCAT CCATGTTGGA ATTTAATCGC GGCCTCGAGC AAGACGTTTC
      TTTAGTCGTA GGTACAACCT TAAATTAGCG CCGGAGCTCG TTCTGCAAAG

6751  CCGTTGAATA TGGCTCATAA CACCCCTTGT ATTACTGTTT ATGTAAGCAG
      GGCAACTTAT ACCGAGTATT GTGGGGAACA TAATGACAAA TACATTCGTC

6801  ACAGTTTTAT TGTTCATGAT GATATATTTT TATCTTGTGC AATGTAACAT
      TGTCAAAATA ACAAGTACTA CTATATAAAA ATAGAACACG TTACATTGTA
```

DraIII

```
6851  CAGAGATTTT GAGACACAAC GTGGCTTTCC CCCCCCCCCC ATTATTGAAG
      GTCTCTAAAA CTCTGTGTTG CACCGAAAGG GGGGGGGGGG TAATAACTTC
.....
6901  CATTTATCAG GGTTATTGTC TCATGAGCGG ATACATATTT GAATGTATTT
      GTAAATAGTC CCAATAACAG AGTACTCGCC TATGTATAAA CTTACATAAA
.....
6951  AGAAAAATAA ACAAATAGGG GTTCCGCGCA CATTTCCCCG AAAAGTGCCA
      TCTTTTTATT TGTTTATCCC CAAGGCGCGT GTAAAGGGGC TTTTCACGGT
.....
7001  CCTGACGTCT AAGAAACCAT TATTATCATG ACATTAACCT ATAAAAATAG
      GGACTGCAGA TTCTTTGGTA ATAATAGTAC TGTAATTGGA TATTTTTATC
.....
7051  GCGTATCACG AGGCCCTTTC GTC
      CGCATAGTGC TCCGGGAAAG CAG
.....
```

SEQUENCE LISTING ID NO: 3

pVR 1012-GP(Z)

General Description

DNA pVR 1012-GP(Z)
Local object
Created: 09/15/98 05:06PM
Last Modification Date: ? (no data)
length: 7285 bp
storage type: Basic
form.: Circular
Comments

Restriction Map

DraIII: 1 site CACNNNGTG / GTGNNNCAC

HindIII: 1 site AAGCTT / TTCGAA

HpaI: 1 site GTTAAC / CAATTG

KasI: 1 site GGCGCC / CCGCGG

NarI: 1 site GGCGCC / CCGCGG

NotI: 1 site GCGGCCGC / CGCCGGCG

PmlI: 1 site CACGTG / GTGCAC

PvuI: 1 site CGATCG / GCTAGC

SacII: 1 site CCGCGG / GGCGCC

XbaI: 1 site TCTAGA / AGATCT

XhoI: 1 site CTCGAG / GAGCTC

EcoRV: 2 sites GATATC / CTATAG

NcoI: 2 sites CCATGG / GGTACC

NdeI: 2 sites CATATG / GTATAC

SphI: 2 sites GCATGC / CGTACG

Functional Map

CDS (4 signals)

CMV IE 5' UT
Start: 886 End: 1129

CMV IE INT
Start: 1130 End: 1860

TbGH
Start: 4302 End: 4854

Kan r
Start: 6350 End: 6972 (Complementary)

Misc_feature (2 signals)

CMV enhancer

Start: 268 End: 885

GP(Z)

Start: 1870 End: 6301

Annotations

```
   1  TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG
      AGCGCGCAAA GCCACTACTG CCACTTTTGG AGACTGTGTA CGTCGAGGGC

  51  GAGACGGTCA CAGCTTGTCT GTAAGCGGAT GCCGGGAGCA GACAAGCCCG
      CTCTGCCAGT GTCGAACAGA CATTCGCCTA CGGCCCTCGT CTGTTCGGGC

 101  TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG TCGGGGCTGG CTTAACTATG
      AGTCCCGCGC AGTCGCCCAC AACCGCCCAC AGCCCCGACC GAATTGATAC

                                        NdeI
 151  CGGCATCAGA GCAGATTGTA CTGAGAGTGC ACCATATGCG GTGTGAAATA
      GCCGTAGTCT CGTCTAACAT GACTCTCACG TGGTATACGC CACACTTTAT

 201  CCGCACAGAT GCGTAAGGAG AAAATACCGC ATCAGATTGG CTATTGGCCA
      GGCGTGTCTA CGCATTCCTC TTTTATGGCG TAGTCTAACC GATAACCGGT

 251  TTGCATACGT TGTATCCATA TCATAATATG TACATTTATA TTGGCTCATG
      AACGTATGCA ACATAGGTAT AGTATTATAC ATGTAAATAT AACCGAGTAC

 301  TCCAACATTA CCGCCATGTT GACATTGATT ATTGACTAGT TATTAATAGT
      AGGTTGTAAT GGCGGTACAA CTGTAACTAA TAACTGATCA ATAATTATCA

 351  AATCAATTAC GGGGTCATTA GTTCATAGCC CATATATGGA GTTCCGCGTT
      TTAGTTAATG CCCCAGTAAT CAAGTATCGG GTATATACCT CAAGGCGCAA

 401  ACATAACTTA CGGTAAATGG CCCGCCTGGC TGACCGCCCA ACGACCCCCG
      TGTATTGAAT GCCATTTACC GGGCGGACCG ACTGGCGGGT TGCTGGGGGC

 451  CCCATTGACG TCAATAATGA CGTATGTTCC CATAGTAACG CCAATAGGGA
      GGGTAACTGC AGTTATTACT GCATACAAGG GTATCATTGC GGTTATCCCT

 501  CTTTCCATTG ACGTCAATGG GTGGAGTATT TACGGTAAAC TGCCCACTTG
      GAAAGGTAAC TGCAGTTACC CACCTCATAA ATGCCATTTG ACGGGTGAAC

                        NdeI
 551  GCAGTACATC AAGTGTATCA TATGCCAAGT ACGCCCCCTA TTGACGTCAA
      CGTCATGTAG TTCACATAGT ATACGGTTCA TGCGGGGGAT AACTGCAGTT

 601  TGACGGTAAA TGGCCCCCCT GGCATTATGC CCAGTACATG ACCTTATGGG
      ACTGCCATTT ACCGGGCGGA CCGTAATACG GGTCATGTAC TGGAATACCC

                                                      NcoI
 651  ACTTTCCTAC TTGGCAGTAC ATCTACGTAT TAGTCATCGC TATTACCATG
      TGAAAGGATG AACCGTCATG TAGATGCATA ATCAGTAGCG ATAATGGTAC

      NcoI
 701  GTGATGCGGT TTTGGCAGTA CATCAATGGG CGTGGATAGC GGTTTGACTC
      CACTACGCCA AAACCGTCAT GTAGTTACCC GCACCTATCG CCAAACTGAG

 751  ACGGCGATTT CCAAGTCTCC ACCCCATTGA CGTCAATGGG AGTTTGTTTT
      TGCCGCTAAA GGTTCAGAGG TGGGGTAACT GCAGTTACCC TCAAACAAAA

 801  GGCACCAAAA TCAACGGGAC TTTCCAAAAT GTCGTAACAA CTCCGCCCCA
      CCGTGGTTTT AGTTGCCCTG AAAGGTTTTA CAGCATTGTT GAGGCGGGGT
```

```
 851  TTGACGCAAA TGGGCGGTAG GCGTGTACGG TCGGAGGTCT ATATAAGCAG
      AACTGCGTTT ACCCGCCATC CGCACATGCC AGCCTCCAGA TATATTCGTC

 901  AGCTCGTTTA GTGAACCGTC AGATCGCCTG GAGACGCCAT CCACGCTGTT
      TCGAGCAAAT CACTTGGCAG TCTAGCGGAC CTCTGCGGTA GGTGCGACAA

                                                  SacII
                                                 ------

 951  TTGACCTCCA TAGAAGACAC CGGGACCGAT CCAGCCTCCG CGGCCGGGAA
      AACTGGAGGT ATCTTCTGTG GCCCTGGCTA GGTCGGAGGC GCCGGCCCTT

1001  CGGTGCATTG GAACGCGGAT TCCCCGTGCC AAGAGTGACG TAAGTACCGC
      GCCACGTAAC CTTGCGCCTA AGGGGCACGG TTCTCACTGC ATTCATGGCG

                                                  SphI
                                                 ------

1051  CTATAGACTC TATAGGCACA CCCCTTTGGC TCTTATGCAT GCTATACTGT
      GATATCTGAG ATATCCGTGT GGGGAAACCG AGAATACGTA CGATATGACA

1101  TTTTGGCTTG GGGCCTATAC ACCCCCGCTT CCTTATGCTA TAGGTGATGG
      AAAACCGAAC CCCGGATATG TGGGGGCGAA GGAATACGAT ATCCACTACC

1151  TATAGCTTAG CCTATAGGTG TGGGTTATTG ACCATTATTG ACCACTCCCC
      ATATCGAATC GGATATCCAC ACCCAATAAC TGGTAATAAC TGGTGAGGGG

1201  TATTGGTGAC GATACTTTCC ATTACTAATC CATAACATGG CTCTTTGCCA
      ATAACCACTG CTATGAAAGG TAATGATTAG GTATTGTACC GAGAAACGGT

1251  CAACTATCTC TATTGGCTAT ATGCCAATAC TCTGTCCTTC AGAGACTGAC
      GTTGATAGAG ATAACCGATA TACGGTTATG AGACAGGAAG TCTCTGACTG

1301  ACGGACTCTC TATTTTTACA GGATGGGGTC CCATTTATTA TTTACAAATT
      TGCCTGAGAC ATAAAAATGT CCTACCCCAG GGTAAATAAT AAATGTTTAA

1351  CACATATACA ACAACGCCGT CCCCCGTGCC CGCAGTTTTT ATTAAACATA
      GTGTATATGT TGTTGCGGCA GGGGGCACGG GCGTCAAAAA TAATTTGTAT

1401  GCGTGGGATC TCCACGCGAA TCTCGGGTAC GTGTTCCGGA CATGGGCTCT
      CGCACCCTAG AGGTGCGCTT AGAGCCCATG CACAAGGCCT GTACCCGAGA

1451  TCTCCGGTAG CGGCGGAGCT TCCACATCCG AGCCCTGGTC CCATGCCTCC
      AGAGGCCATC GCCGCCTCGA AGGTGTAGGC TCGGGACCAG GGTACGGAGG

1501  AGCGGCTCAT GCTCGCTCGG CAGCTCCTTG CTCCTAACAG TGGAGGCCAG
      TCGCCGAGTA CGAGCGAGCC GTCGAGGAAC GAGGATTGTC ACCTCCGGTC

1551  ACTTAGGCAC AGCACAATGC CCACCACCAC CAGTGTGCCG CACAAGGCCG
      TGAATCCGTG TCGTGTTACG GGTGGTGGTG GTCACACGGC GTGTTCCGGC

1601  TGGCGGTAGG GTATGTGTCT GAAAATGAGC GTGGAGATTG GGCTCGCACG
      ACCGCCATCC CATACACAGA CTTTTACTCG CACCTCTAAC CCGAGCGTGC

1651  GCTGACGCAG ATGGAAGACT TAAGGCAGCG GCAGAAGAAG ATGCAGGCAG
      CGACTGCGTC TACCTTCTGA ATTCCGTCGC CGTCTTCTTC TACGTCCGTC

1701  CTGAGTTCTT GTATTCTGAT AAGAGTCAGA GGTAACTCCC GTTGCGGTGC
      GACTCAAGAA CATAAGACTA TTCTCAGTCT CCATTGAGGG CAACGCCACG
```

```
     HpaI
1751 TGTTAACGGT GGAGGGCAGT GTAGTCTGAG CAGTACTCGT TGCTGCCGCG
     ACAATTGCCA CCTCCCGTCA CATCAGACTC GTCATGAGCA ACGACGGCGC

                                                   NcoI
1801 CGCGCCACCA GACATAATAG CTGACAGACT AACAGACTGT TCCTTTCCAT
     GCGCGGTGGT CTGTATTATC GACTGTCTGA TTGTCTGACA AGGAAAGGTA

     NcoI                             PmlI       EcoRVNotI
1851 GGGTCTTTTC TGCAGTCACC GTCGTCGACA CGTGTGATCA GATATCGCGG
     CCCAGAAAAG ACGTCAGTGG CAGCAGCTGT GCACACTAGT CTATAGCGCC

                 NarI
     NotI XbaI   KasI
1901 CCGCTCTAGA CCAGGCGCCT GGATCGATCC GCGATGAAGA TTAAGCCGAC
     GGCGAGATCT GGTCCGCGGA CCTAGCTAGG CGCTACTTCT AATTCGGCTG

1951 AGTGAGCGTA ATCTTCATCT CTCTTAGATT ATTTGTTTTC CAGAGTAGGG
     TCACTCGCAT TAGAAGTAGA GAGAATCTAA TAAACAAAAG GTCTCATCCC

2001 GTCGTCAGGT CCTTTTCAAT CGTGTAACCA AAATAAACTC CACTAGAAGG
     CAGCAGTCCA GGAAAAGTTA GCACATTGGT TTTATTTGAG GTGATCTTCC

2051 ATATTGTGGG GCAACAACAC AATGGGCGTT ACAGGAATAT TGCAGTTACC
     TATAACACCC CGTTGTTGTG TTACCCGCAA TGTCCTTATA ACGTCAATGG

2101 TCGTGATCGA TTCAAGAGGA CATCATTCTT TCTTTGGGTA ATTATCCTTT
     AGCACTAGCT AAGTTCTCCT GTAGTAAGAA AGAAACCCAT TAATAGGAAA

2151 TCCAAAGAAC ATTTTCCATC CCACTTGGAG TCATCCACAA TAGCACATTA
     AGGTTTCTTG TAAAAGGTAG GGTGAACCTC AGTAGGTGTT ATCGTGTAAT

2201 CAGCTTAGTG ATGTCGACAA ACTAGTTTGT CGTGACAAAC TGTCATCCAC
     GTCCAATCAC TACAGCTGTT TGATCAAACA GCACTGTTTG ACAGTAGGTG

2251 AAATCAATTG AGATCAGTTG GACTGAATCT CGAAGGGAAT GGAGTGGCAA
     TTTAGTTAAC TCTAGTCAAC CTGACTTAGA GCTTCCCTTA CCTCACCGTT

2301 CTGACGTGCC ATCTGCAACT AAAAGATGGG GCTTCAGGTC CGGTGTCCCA
     GACTGCACGG TAGACGTTGA TTTTCTACCC CGAAGTCCAG GCCACAGGGT

2351 CCAAAGGTGG TCAATTATGA AGCTGGTGAA TGGGCTGAAA ACTGCTACAA
     GGTTTCCACC AGTTAATACT TCGACCACTT ACCCGACTTT TGACGATGTT

2401 TCTTGAAATC AAAAAACCTG ACGGCAGTGA GTGTCTACCA GCAGCGCCAG
     AGAACTTTAG TTTTTTGGAC TGCCCTCACT CACAGATGGT CGTCGCGGTC

2451 ACGGGATTCG GGGCTTCCCC CGGTGCCGGT ATGTGCACAA AGTATCAGGA
     TGCCCTAAGC CCCGAAGGGG GCCACGGCCA TACACGTGTT TCATAGTCCT

2501 ACGGGACCGT GTGCCGGAGA CTTTGCCTTC CATAAAGAGG GTGCTTTCTT
     TGCCCTGGCA CACGGCCTCT GAAACGGAAG GTATTTCTCC CACGAAAGAA
```

```
2551  CCTGTATGAT CGACTTGCTT CCACAGTTAT CTACCGAGGA ACGACTTTCG
      GGACATACTA GCTGAACGAA GGTGTCAATA GATGGCTCCT TGCTGAAAGC
.....
2601  CTGAAGGTGT CGTTGCATTT CTGATACTGC CCCAAGCTAA GAAGGACTTC
      GACTTCCACA GCAACGTAAA GACTATGACG GGGTTCGATT CTTCCTGAAG
.....
2651  TTCAGCTCAC ACCCCTTGAG AGAGCCGGTC AATGCAACGG AGGACCCGTC
      AAGTCGAGTG TGGGGAACTC TCTCGGCCAG TTACGTTGCC TCCTGGGCAG
.....
                                    EcoRV
                                   ------
2701  TAGTGGCTAC TATTCTACCA CAATTAGATA TCAGGCTACC GGTTTTGGAA
      ATCACCGATG ATAAGATGGT GTTAATCTAT AGTCCGATGG CCAAAACCTT
.....
2751  CCAATGAGAC AGAGTACTTG TTCGAGGTTG ACAATTTGAC CTACGTCCAA
      GGTTACTCTG TCTCATGAAC AAGCTCCAAC TGTTAAACTG GATGCAGGTT
.....
2801  CTTGAATCAA GATTCACACC ACAGTTTCTG CTCCAGCTGA ATGAGACAAT
      GAACTTAGTT CTAAGTGTGG TGTCAAAGAC GAGGTCGACT TACTCTGTTA
.....
2851  ATATACAAGT GGGAAAAGGA GCAATACCAC GGGAAAACTA ATTTGGAAGG
      TATATGTTCA CCCTTTTCCT CGTTATGGTG CCCTTTTGAT TAAACCTTCC
.....
2901  TCAACCCCGA AATTGATACA ACAATCGGGG AGTGGGCCTT CTGGGAAACT
      AGTTGGGGCT TTAACTATGT TGTTAGCCCC TCACCCGGAA GACCCTTTGA
.....
2951  AAAAAAACC TCACTAGAAA AATTCGCAGT GAAGAGTTGT CTTTCACAGT
      TTTTTTTGG AGTGATCTTT TTAAGCGTCA CTTCTCAACA GAAAGTGTCA
.....
3001  TGTATCAAAC GGAGCCAAAA ACATCAGTGG TCAGACTCCG GCGCGAACTT
      ACATAGTTTG CCTCGGTTTT TGTAGTCACC AGTCTCAGGC CGCGCTTGAA
.....
3051  CTTCCGACCC AGGGACCAAC ACAACAACTG AAGACCACAA AATCATGGCT
      GAAGGCTGGG TCCCTGGTTG TGTTGTTGAC TTCTGGTGTT TTAGTACCGA
.....
3101  TCAGAAAATT CCTCCGCAAT GGTTCAAGTG CACAGTCAAG GAAGGGAAGC
      AGTCTTTTAA GGAGACGTTA CCAAGTTCAC GTGTCAGTTC CTTCCCTTCG
.....
3151  TGCAGTGTCG CATCTAACAA CCCTTGCCAC AATCTCCACG AGTCCCCAAT
      ACGTCACAGC GTAGATTGTT GGGAACGGTG TTAGAGGTGC TCAGGGGTTA
.....
3201  CCCTCACAAC CAAACCAGGT CCGGACAACA GCACCCATAA TACACCCGTG
      GGGAGTGTTG GTTTGGTCCA GGCCTGTTGT CGTGGGTATT ATGTGGGCAC
.....
3251  TATAAACTTG ACATCTCTGA GGCAACTCAA GTTGAACAAC ATCACCGCAG
      ATATTTGAAC TGTAGAGACT CCGTTGAGTT CAACTTGTTG TAGTGGCGTC
.....
3301  AACAGACAAC GACAGCACAG CCTCCGACAC TCCCTCTGCC ACGACCGCAG
      TTGTCTGTTG CTGTCGTGTC GGAGGCTGTG AGGGAGACGG TGCTGGCGTC
.....
3351  CCGGACCCCC AAAAGCAGAG AACACCAACA CGAGCAAGAG CACTGACTTC
      GGCCTGGGGG TTTTCGTCTC TTGTGGTTGT GCTCGTTCTC GTGACTGAAG
.....
3401  CTGGACCCCG CCACCACAAC AAGTCCCCAA AACCACAGCG AGACCGCTGG
      GACCTGGGGC GGTGGTGTTG TTCAGGGGTT TTGGTGTCGC TCTGGCGACC
.....
```

```
3451  CAACAACAAC ACTCATCACC AAGATACCGG AGAAGAGAGT GCCAGCAGCG
      GTTGTTGTTG TGAGTAGTGG TTCTATGGCC TCTTCTCTCA CGGTCGTCGC
.....
3501  GGAAGCTAGG CTTAATTACC AATACTATTG CTGGAGTCGC ACGACTGATC
      CCTTCGATCC GAATTAATGG TTATGATAAC GACCTCAGCG TCCTGACTAG
.....
3551  ACAGGCGGGA GAAGAACTCG AAGAGAAGCA ATTGTCAATG CTCAACCCAA
      TGTCCGCCCT CTTCTTGAGC TTCTCTTCGT TAACAGTTAC GAGTTGGGTT
.....
3601  ATGCAACCCT AATTTACATT ACTGGACTAC TCAGGATGAA GGTGCTGCAA
      TACGTTGGGA TTAAATGTAA TGACCTGATG AGTCCTACTT CCACGACGTT
.....
3651  TCGGACTGGC CTGGATACCA TATTTCGGGC CAGCAGCCGA GGGAATTTAC
      AGCCTGACCG GACCTATGGT ATAAAGCCCG GTCGTCGGCT CCCTTAAATG
.....
3701  ATAGAGGGGC TAATGCACAA TCAAGATGGT CTAATCTGTG GGTTGAGACA
      TATCTCCCCG ATTACGTGTT AGTTCTACCA AATTAGACAC CCAACTCTGT
.....
3751  GCTGGCCAAC GAGACGACTC AAGCTCTTCA ACTGTTCCTG AGAGCCACAA
      CGACCGGTTG CTCTGCTGAG TTCGAGAAGT TGACAAGGAC TCTCGGTGTT
.....
3801  CTGAGCTACG CACCTTTTCA ATCCTCAACC GTAAGGCAAT TGATTCCTTG
      GACTCGATGC GTGGAAAAGT TAGGAGTTGG CATTCCGTTA ACTAAAGAAC
.....
3851  CTGCAGCGAT GGGGCGGCAC ATGCCACATT CTGGGACCGG ACTGCTGTAT
      GACGTCGCTA CCCCGCCGTG TACGGTGTAA GACCCTGGCC TGACGACATA
.....
3901  CGAACCACAT GATTGGACCA AGAACATAAC AGACAAAATC GATCAGATTA
      GCTTGGTGTA CTAACCTGGT TCTTGTATTG TCTGTTTTAG CTAGTCTAAT
.....
3951  TTCATGATTT TGTTGATAAA ACCCTTCCGG ACCAGGGGGA CAATGACAAT
      AAGTACTAAA ACAACTATTT TGGGAAGGCC TGGTCCCCCT GTTACTGTTA
.....
4001  TGGTCGACAG GATGGAGACA ATGGATACCG GCAGGTATTG GAGTTACAGG
      ACCACCTGTC CTACCTCTGT TACCTATGGC CGTCCATAAC CTCAATGTCC
.....
4051  CGTTATAATT GCAGTTATCG CTTTATTCTG TATATGCAAA TTTGTCTTTT
      GCAATATTAA CGTCAATAGC GAAATAAGAC ATATACGTTT AAACAGAAAA
.....
4101  AGTTTTTCTT CAGATTGCTT CATGGAAAAG CTCAGCCTCA AATCAATGAA
      TCAAAAAGAA GTCTAACGAA GTACCTTTTC GAGTCGGAGT TTAGTTACTT
.....
4151  ACCAGGATTT AATTATATGG ATTACTTGAA TCTAAGATTA CTTGACAAAT
      TGGTCCTAAA TTAATATACC TAATGAACTT AGATTCTAAT GAACTGTTTA
.....
4201  GATAATATAA TACACTGGAG CTTTAAACAT AGCCAATGTG ATTCTAACTC
      CTATTATATT ATGTGACCTC GAAATTTGTA TCGGTTACAC TAAGATTGAG
.....
4251  CTTTAAACTC ACAGTTAATC ATAAACAAGG TTTGGTACCG AGCTCGAATT
      GAAATTTGAG TGTCAATTAG TATTTGTTCC AAACCATGGC TCGAGCTTAA
.....
4301  ATCTGCTGTG CCTTCTAGTT GCCAGCCATC TGTTGTTTGC CCCTCCCCCG
      TAGACGACAC GGAAGATCAA CGGTCGGTAG ACAACAAACG GGGAGGGGGC
.....
4351  TGCCTTCCTT GACCCTGGAA GGTGCCACTC CCACTGTCCT TTCCTAATAA
      ACGGAAGGAA CTGGGACCTT CCACGGTGAG GGTGACAGGA AAGGATTATT
.....
```

```
4401  AATGAGGAAA TTGCATCGCA TTGTCTGAGT AGGTGTCATT CTATTCTGGG
      TTACTCCTTT AACGTAGCGT AACAGACTCA TCCACAGTAA GATAAGACCC
.....
4451  GGGTGGGGTG GGGCAGCACA GCAAGGGGGA GGATTGGGAA GACAATAGCA
      CCCACCCCAC CCCGTCGTGT CGTTCCCCCT CCTAACCCTT CTGTTATCGT
.....
        SphI
      ------
4501  GGCATGCTGG GGATGCGGTG GGCTCTATGG GTACCCAGGT GCTGAAGAAT
      CCGTACGACC CCTACGCCAC CCGAGATACC CATGGGTCCA CGACTTCTTA
.....
4551  TGACCCGGTT CCTCCTGGGC CAGAAAGAAG CAGGCACATC CCCTTCTCTG
      ACTGGGCCAA GGAGGACCCG GTCTTTCTTC GTCCGTGTAG GGGAAGAGAC
.....
4601  TGACACACCC TGTCCACGCC CCTGGTTCTT AGTTCCAGCC CCACTCATAG
      ACTGTGTGGG ACAGGTGCGG GGACCAAGAA TCAAGGTCGG GGTGAGTATC
.....
4651  GACACTCATA GCTCAGGAGG GCTCCGCCTT CAATCCCACC CGCTAAAGTA
      CTGTGAGTAT CGAGTCCTCC CGAGGCGGAA GTTAGGGTGG GCGATTTCAT
.....
4701  CTTGGAGCGG TCTCTCCCTC CCTCATCAGC CCACCAAACC AAACCTAGCC
      GAACCTCGCC AGAGAGGGAG GGAGTAGTCG GGTGGTTTGG TTTGGATCGG
.....
4751  TCCAACAGTG GCAAGAAATT AAAGCAAGAT AGGCTATTAA GTGCAGAGGG
      AGGTTGTCAC CGTTCTTTAA TTTCGTTCTA TCCGATAATT CACGTCTCCC
.....
4801  AGAGAAAATG CCTCCAACAT GTGAGGAAGT AATGAGAGAA ATCATAGAAT
      TCTCTTTTAC GGAGGTTGTA CACTCCTTCA TTACTCTCTT TAGTATCTTA
.....
4851  TTCTTCCGCT TCCTCGCTCA CTGACTCGCT GCGCTCGGTC GTTCGGCTGC
      AAGAAGGCGA AGGAGCGAGT GACTGAGCGA CGCGAGCCAG CAAGCCGACG
.....
4901  GGCGAGCGGT ATCAGCTCAC TCAAAGGCGG TAATACGGTT ATCCACAGAA
      CCGCTCGCCA TAGTCGAGTG AGTTTCCGCC ATTATGCCAA TAGGTGTCTT
.....
4951  TCAGGGGATA ACGCAGGAAA GAACATGTGA GCAAAAGGCC AGCAAAAGGC
      AGTCCCCTAT TGCGTCCTTT CTTGTACACT CGTTTTCCGG TCGTTTTCCG
.....
5001  CAGGAACCGT AAAAAGGCCG CGTTGCTGGC GTTTTTCCAT AGGCTCCGCC
      GTCCTTGGCA TTTTTCCGGC GCAACGACCG CAAAAAGGTA TCCGAGGCGG
.....
5051  CCCCTGACGA GCATCACAAA AATCGACGCT CAAGTCAGAG GTGGCGAAAC
      GGGGACTGCT CGTAGTGTTT TTAGCTGCGA GTTCAGTCTC CACCGCTTTG
.....
5101  CCGACAGGAC TATAAAGATA CCAGGCGTTT CCCCCTGGAA GCTCCCTCGT
      GGCTGTCCTG ATATTTCTAT GGTCCGCAAA GGGGGACCTT CGAGGGAGCA
.....
5151  GCGCTCTCCT GTTCCGACCC TGCCGCTTAC CGGATACCTG TCCGCCTTTC
      CGCGAGAGGA CAAGGCTGGG ACGGCGAATG GCCTATGGAC AGGCGGAAAG
.....
5201  TCCCTTCGGG AAGCGTGGCG CTTTCTCAAT GCTCACGCTG TAGGTATCTC
      AGGGAAGCCC TTCGCACCGC GAAAGAGTTA CGAGTGCGAC ATCCATAGAG
.....
5251  AGTTCGGTGT AGGTCGTTCG CTCCAAGCTG GGCTGTGTGC ACGAACCCCC
      TCAAGCCACA TCCAGCAAGC GAGGTTCGAC CCGACACACG TGCTTGGGGG
.....
```

```
5301  CGTTCACCCC GACCGCTGCG CCTTATCCGG TAACTATCGT CTTGAGTCCA
      GCAAGTCGGG CTGGCGACGC GGAATAGGCC ATTGATAGCA GAACTCAGGT
.....
5351  ACCCGGTAAG ACACGACTTA TCGCCACTGG CAGCAGCCAC TGGTAACAGG
      TGGGCCATTC TGTGCTGAAT AGCGGTGACC GTCGTCGGTG ACCATTGTCC
.....
5401  ATTAGCAGAG CGAGGTATGT AGGCGGTGCT ACAGAGTTCT TGAAGTGGTG
      TAATCGTCTC GCTCCATACA TCCGCCACGA TGTCTCAAGA ACTTCACCAC
.....
5451  GCCTAACTAC GGCTACACTA GAAGGACAGT ATTTGGTATC TGCGCTCTGC
      CGGATTGATG CCGATGTGAT CTTCCTGTCA TAAACCATAG ACGCGAGACG
.....
5501  TGAAGCCAGT TACCTTCGGA AAAAGAGTTG GTAGCTCTTG ATCCGGCAAA
      ACTTCGGTCA ATGGAAGCCT TTTTCTCAAC CATCGAGAAC TAGGCCGTTT
.....
5551  CAAACCACCG CTGGTAGCGG TGGTTTTTTT GTTTGCAAGC AGCAGATTAC
      GTTTGGTGGC GACCATCGCC ACCAAAAAAA CAAACGTTCG TCGTCTAATG
.....
5601  GCGCAGAAAA AAAGGATCTC AAGAAGATCC TTTGATCTTT TCTACGGGGT
      CGCGTCTTTT TTTCCTAGAG TTCTTCTAGG AAACTAGAAA AGATGCCCCA
.....
5651  CTGACGCTCA GTGGAACGAA AACTCACGTT AAGGGATTTT GGTCATGAGA
      GACTGCGAGT CACCTTGCTT TTGAGTGCAA TTCCCTAAAA CCAGTACTCT
.....
5701  TTATCAAAAA GGATCTTCAC CTAGATCCTT TTAAATTAAA AATGAAGTTT
      AATAGTTTTT CCTAGAAGTG GATCTAGGAA AATTTAATTT TTACTTCAAA
.....
5751  TAAATCAATC TAAAGTATAT ATGAGTAAAC TTGGTCTGAC AGTTACCAAT
      ATTTAGTTAG ATTTCATATA TACTCATTTG AACCAGACTG TCAATGGTTA
.....
5801  GCTTAATCAG TGAGGCACCT ATCTCAGCGA TCTGTCTATT TCGTTCATCC
      CGAATTAGTC ACTCCGTGGA TAGAGTCGCT AGACAGATAA AGCAAGTAGG
.....
5851  ATAGTTGCCT GACTCCGGGG GGGGGGGGCG CTGAGGTCTG CCTCGTGAAG
      TATCAACGGA CTGAGGCCCC CCCCCCCCGC GACTCCAGAC GGAGCACTTC
.....
5901  AAGGTGTTGC TGACTCATAC CAGGCCTGAA TCGCCCCATC ATCCAGCCAG
      TTCCACAACG ACTGAGTATG GTCCGGACTT AGCGGGGTAG TAGGTCGGTC
.....
5951  AAAGTGAGGG AGCCACGGTT GATGAGAGCT TTGTTGTAGG TGGACCAGTT
      TTTCACTCCC TCGGTGCCAA CTACTCTCGA AACAACATCC ACCTGGTCAA
.....
6001  GGTGATTTTG AACTTTTGCT TTGCCACGGA ACGGTCTGCG TTGTCGGGAA
      CCACTAAAAC TTGAAAACGA AACGGTGCCT TCCCAGACGC AACAGCCCTT
.....
6051  GATGCGTGAT CTGATCCTTC AACTCAGCAA AAGTTCGATT TATTCAACAA
      CTACGCACTA GACTAGGAAG TTGAGTCGTT TTCAAGCTAA ATAAGTTGTT
.....
6101  AGCCGCCGTC CCGTCAAGTC AGCGTAATGC TCTGCCAGTG TTACAACCAA
      TCGGCGGCAG GGCAGTTCAG TCGCATTACG AGACGGTCAC AATGTTGGTT
.....
6151  TTAACCAATT CTGATTAGAA AAACTCATCG AGCATCAAAT GAAACTGCAA
      AATTGGTTAA GACTAATCTT TTTGAGTAGC TCGTAGTTTA CTTTGACGTT
.....
6201  TTTATTCATA TCAGGATTAT CAATACCATA TTTTTGAAAA AGCCGTTTCT
      AAATAAGTAT AGTCCTAATA GTTATGGTAT AAAAACTTTT TCGGCAAAGA
.....
```

```
6251  GTAATGAAGG AGAAAACTCA CCGAGGCAGT TCCATAGGAT GGCAAGATCC
      CATTACTTCC TCTTTTGAGT GGCTCCGTCA AGGTATCCTA CCGTTCTAGG
.....

6301  TGGTATCGGT CTGCGATTCC GACTCGTCCA ACATCAATAC AACCTATTAA
      ACCATAGCCA GACGCTAAGG CTGAGCAGGT TGTAGTTATG TTGGATAATT
.....

6351  TTTCCCCTCG TCAAAAATAA GGTTATCAAG TGAGAAATCA CCATGAGTGA
      AAAGGGGAGC AGTTTTTATT CCAATAGTTC ACTCTTTAGT GGTACTCACT
.....

                              HindIII
                              ========

6401  CGACTGAATC CGGTGAGAAT GGCAAAAGCT TATGCATTTC TTTCCAGACT
      GCTGACTTAG GCCACTCTTA CCGTTTTCGA ATACGTAAAG AAAGGTCTGA
.....

6451  TGTTCAACAG GCCAGCCATT ACGCTCGTCA TCAAAATCAC TCGCATCAAC
      ACAAGTTGTC CGGTCGGTAA TGCGAGCAGT AGTTTTAGTG AGCGTAGTTG
.....

                                                      PvuI
                                                      ====

6501  CAAACCGTTA TTCATTCGTG ATTGCGCCTG AGCGAGACGA AATACGCGAT
      GTTTGGCAAT AAGTAAGCAC TAACGCGGAC TCGCTCTGCT TTATGCGCTA
.....

      PvuI
      ====

6551  CGCTGTTAAA AGGACAATTA CAAACAGGAA TCGAATGCAA CCGGCGCAGG
      GCGACAATTT TCCTGTTAAT GTTTGTCCTT AGCTTACGTT GGCCGCGTCC
.....

6601  AACACTGCCA GCGCATCAAC AATATTTTCA CCTGAATCAG GATATTCTTC
      TTGTGACGGT CGCGTAGTTG TTATAAAAGT GGACTTAGTC CTATAAGAAG
.....

6651  TAATACCTGG AATGCTGTTT TCCCGGGGAT CGCAGTGGTG AGTAACCATG
      ATTATGGACC TTACGACAAA AGGGCCCCTA GCGTCACCAC TCATTGGTAC
.....

6701  CATCATCAGG AGTACGGATA AAATGCTTGA TGGTCGGAAG AGGCATAAAT
      GTAGTAGTCC TCATGCCTAT TTTACGAACT ACCAGCCTTC TCCGTATTTA
.....

6751  TCCGTCAGCC AGTTTAGTCT GACCATCTCA TCTGTAACAT CATTGGCAAC
      AGGCAGTCGG TCAAATCAGA CTGGTAGAGT AGACATTGTA GTAACCGTTG
.....

6801  GCTACCTTTG CCATGTTTCA GAAACAACTC TGGCGCATCG GGCTTCCCAT
      CGATGGAAAC GGTACAAAGT CTTTGTTGAG ACCGCGTAGC CCGAAGGGTA
.....

6851  ACAATCGATA GATTGTCGCA CCTGATTGCC CGACATTATC GCGAGCCCAT
      TGTTAGCTAT CTAACAGCGT GGACTAACGG GCTGTAATAG CGCTCGGGTA
.....

                                                      XhoI
                                                      ====

6901  TTATACCCAT ATAAATCAGC ATCCATGTTG GAATTTAATC GCGGCCTCGA
      AATATGGGTA TATTTAGTCG TAGGTACAAC CTTAAATTAG CGCCGGAGCT
.....

      XhoI
      ====

6951  GCAAGACGTT TCCCGTTGAA TATGGCTCAT AACACCCCTT GTATTACTGT
      CGTTCTGCAA AGGGCAACTT ATACCGAGTA TTGTGGGGAA CATAATGACA
.....

7001  TTATGTAAGC AGACAGTTTT ATTGTTCATG ATGATATATT TTTATCTTGT
      AATACATTCG TCTGTCAAAA TAACAAGTAC TACTATATAA AAATAGAACA
.....
```

DraIII

```
7051  GCAATGTAAC ATCAGAGATT TTGAGACACA ACGTGGCTTT CCCCCCCCCC
      CGTTACATTG TAGTCTCTAA AACTCTGTGT TGCACCGAAA GGGGGGGGG
.....
7101  CCATTATTGA AGCATTTATC AGGGTTATTG TCTCATGAGC GGATACATAT
      GGTAATAACT TCGTAAATAG TCCCAATAAC AGAGTACTCG CCTATGTATA
.....
7151  TTGAATGTAT TTAGAAAAAT AAACAAATAG GGGTTCCGCG CACATTTCCC
      AACTTACATA AATCTTTTTA TTTGTTTATC CCCAAGGCGC GTGTAAAGGG
.....
7201  CGAAAAGTGC CACCTGACGT CTAAGAAACC ATTATTATCA TGACATTAAC
      GCTTTTCACG GTGGACTGCA GATTCTTTGG TAATAATAGT ACTGTAATTG
.....
7251  CTATAAAAAT AGGCGTATCA CGAGGCCCTT TCGTC
      GATATTTTTA TCCGCATAGT GCTCCGGGAA AGCAG
.....
```

pVR 1012-SGP(Z)

General Description

DNA pVR 1012-SGP(Z)
Local object
Created: 09/14/98 04:29PM
Last Modified: 09/15/98 04:50PM
length: 7272 bp
storage type: Basic
form: Circular
Comments

Restriction Map

DraIII: 1 site CACNNNGTG / GTGNNNCAC

HindIII: 1 site AAGCTT / TTCGAA

HpaI: 1 site GTTAAC / CAATTG

KpnI: 1 site GGTACC / CCATGG

NotI: 1 site GCGGCCGC / CGCCGGCG

PmlI: 1 site CACGTG / GTGCAC

PvuI: 1 site CGATCG / GCTAGC

SacII: 1 site CCGCGG / GGCGCC

XbaI: 1 site TCTAGA / AGATCT

XhoI: 1 site CTCGAG / GAGCTC

EcoRV: 2 sites GATATC / CTATAG

NcoI: 2 sites CCATGG / GGTACC

NdeI: 2 sites CATATG / GTATAC

SphI: 2 sites GCATGC / CGTACG

Functional Map

CDS (4 signals)

CMV IE 5' UT
Start: 886 End: 1129

CMV IE INT
Start: 1130 End: 1840

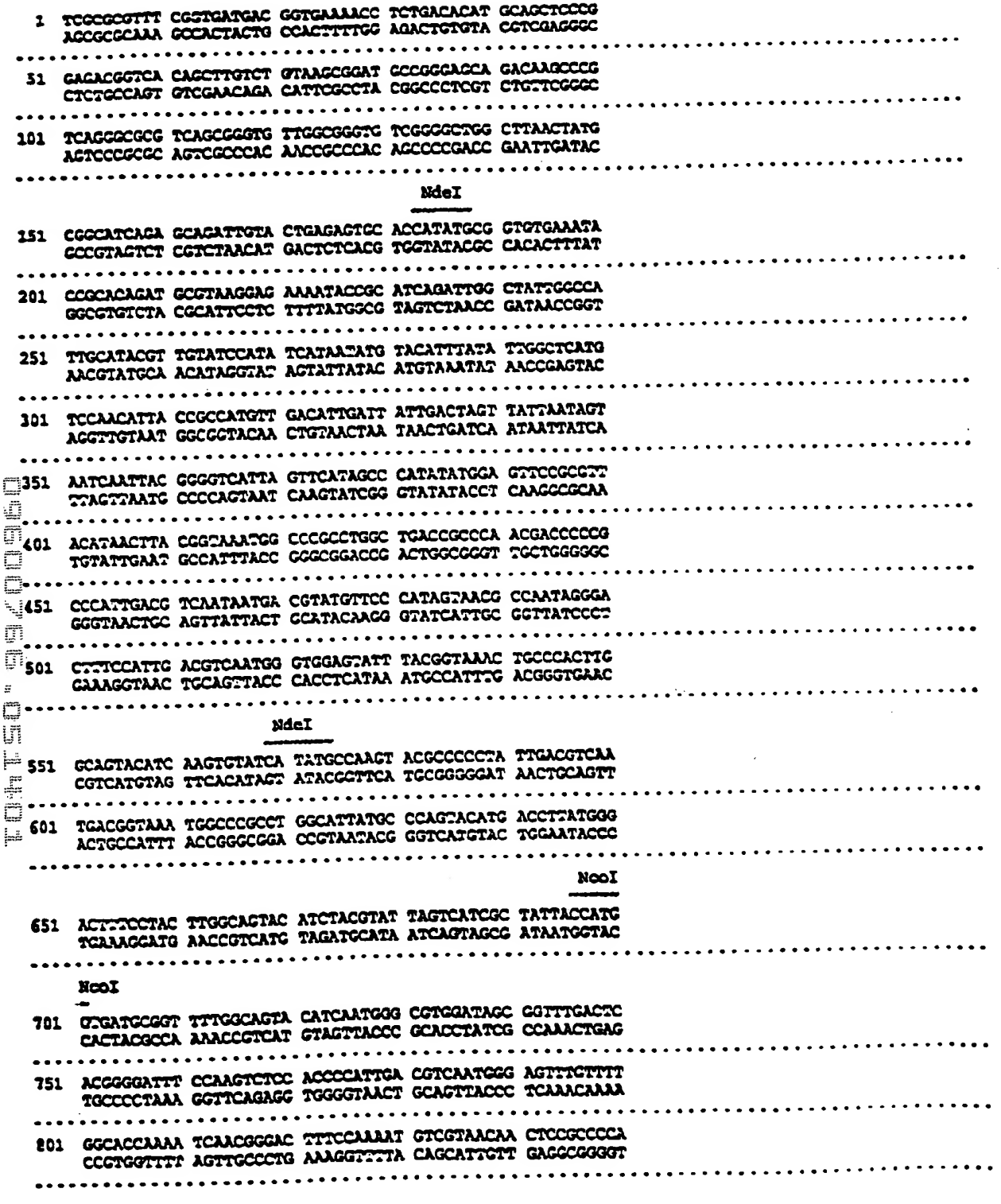
TbGH
Start: 4289 End: 4861

Kanr
Start: 6337 End: 6959 (Complementary)

Misc_feature (2 signals)

CMV enhancer
Start: 248 End: 885
SGP(Z)
Start: 1870 End: 4288

Annotations


NdeI
NdeI
NcoI
NcoI

```
 651  TTGACGCAAA TGGGCGGTAG GCGTGTACGG TGGGAGGTCT ATATAAGCAG
      AACTGCGTTT ACCCGCCATC CGCACATGCC ACCCTCCAGA TATATTCGTC
 ..........................................................................

 901  AGCTCGTTTA GTGAACCGTC AGATCGCCTG GAGACGCCAT CCACGCTGTT
      TCGAGCAAAT CACTTGGCAG TCTAGCGGAC CTCTGCGGTA GGTGCGACAA
 ..........................................................................

                                                  SacII
                                                  ~~~~~~

 951  TTGACCTCCA TAGAAGACAC CGGGACCGAT CCAGCCTCCG CGGCCGGGAA
      AACTGGAGGT ATCTTCTGTG GCCCTGGCTA GGTCGGAGGC GCCGGCCCTT
 ..........................................................................

1001  CGGTGCATTG GAACGCGGAT TCCCCGTGCC AAGAGTGACG TAAGTACCGC
      GCCACGTAAC CTTGCGCCTA AGGGGCACGG TTCTCACTGC ATTCATGGCG
 ..........................................................................

                                                  SphI
                                                  ~~~~~~

1051  CTATAGACTC TATAGGCACA CCCCTTTGGC TCTTATGCAT GCTATACTGT
      GATATCTGAG ATATCCGTGT GGGGAAACCG AGAATACGTA CGATATGACA
 ..........................................................................

1101  TTTTGGCTTG GGGCCTATAC ACCCCCGCTT CCTTATGCTA TAGGTGATGG
      AAAACCGAAC CCCGGATATG TGGGGGCGAA GGAATACGAT ATCCACTACC
 ..........................................................................

1151  TATAGCTTAG CCTATAGGTG TGGGTTATTG ACCATTATTG ACCACTCCCC
      ATATCGAATC GGATATCCAC ACCCAATAAC TGGTAATAAC TGGTGAGGGG
 ..........................................................................

1201  TATTGGTGAC GATACTTTCC ATTACTAATC CATAACATGG CTCTTTGCCA
      ATAACCACTG CTATGAAAGG TAATGATTAG GTATTGTACC GAGAAACGGT
 ..........................................................................

1251  CAACTATCTC TATTGGCTAT ATGCCAATAC TCTGTCCTTC AGAGACTGAC
      GTTGATAGAG ATAACCGATA TACGGTTATG AGACAGGAAG TCTCTGACTG
 ..........................................................................

1301  ACGGACTCTG TATTTTTACA GGATGGGGTC CCATTTATTA TTTACAAATT
      TGCCTGAGAC ATAAAAATGT CCTACCCCAG GGTAAATAAT AAATGTTTAA
 ..........................................................................

1351  CACATATACA ACAACGCCGT CCCCCGTGCC CGCAGTTTTT ATTAAACATA
      GTGTATATGT TGTTGCGGCA GGGGGCACGG GCGTCAAAAA TAATTTGTAT
 ..........................................................................

1401  GCGTGGGATC TCCACGCGAA TCTCGGGTAC GTGTTCCGGA CATGGGCTCT
      CGCACCCTAG AGGTGCGCTT AGAGCCCATG CACAAGGCCT GTACCCGAGA
 ..........................................................................

1451  TCTCCGGTAG CGGCGGAGCT TCCACATCCG AGCCCTGGTC CCATGCCTCC
      AGAGGCCATC GCCGCCTCGA AGGTGTAGGC TCGGGACCAG GGTACGGAGG
 ..........................................................................

1501  AGCGGCTCAT GGTCGCTCGG CAGCTCCTTG CTCCTAACAG TGGAGGCCAG
      TCGCCGAGTA CCAGCGAGCC GTCGAGGAAC GAGGATTGTC ACCTCCGGTC
 ..........................................................................

1551  ACTTAGGCAC AGCACAATGC CCACCACCAC CAGTGTGCCG CACAAGGCCG
      TGAATCCGTG TCGTGTTACG GGTGGTGGTG GTCACACGGC GTGTTCCGGC
 ..........................................................................

1601  TGGCGGTAGG GTATGTGTCT GAAAATGAGC GTGGAGATTG GGCTCGCACG
      ACCGCCATCC CATACACACA CTTTTACTCG CACCTCTAAC CCGAGCGTGC
 ..........................................................................

1651  GCTGACGCAG ATGGAAGACT TAAGGCAGCG GCAGAAGAAG ATGCAGGCAG
      CGACTGCGTC TACCTTCTGA ATTCCGTCGC CGTCTTCTTC TACGTCCGTC
 ..........................................................................

1701  CTGAGTTGTT GTATTCTGAT AAGAGTCAGA GGTAACTCCC GTTGCGGTGC
      GACTCAACAA CATAAGACTA TTCTCAGTCT CCATTGAGGG CAACGCCACG
 ..........................................................................
```

```
      HpaI
1751  TGTTAACGGT GGAGGGCAGT GTAGTCTGAG CAGTACTCGT TGCTGCCGCG
      ACAATTGCCA CCTCCCGTCA CATCAGACTC GTCATGAGCA ACGACGGCGC

                                                        NcoI
1801  CGCCCCACCA GACATAATAG CTGACAGACT AACAGACTGT TCCTTTCCAT
      GCGCGGTGGT CTGTATTATC GACTGTCTGA TTGTCTGACA AGGAAAGGTA

      NcoI                                PmlI        EcoRV NotI
1851  GGGTCTTTTC TGCAGTCACC GTCGTCGACA CGTCTGATCA GATATCGCGG
      CCCAGAAAAG ACGTCAGTGG CAGCAGCTGT GCACACTAGT CTATAGCGCC

      NotI XbaI
1901  CCGCTCTAGA CCAGGCGCCT GGATCGAATT GATGAAGATT AAGCCGACAG
      GGCGAGATCT GGTCCGCGGA CCTAGCTTAA CTACTTCTAA TTCGGCTGTC

1951  TGAGCGTAAT CTTCATCTCT CTTAGATTAT TTGTTTTCCA GAGTAGGGGT
      ACTCGCATTA GAAGTAGAGA GAATCTAATA AACAAAAGGT CTCATCCCCA

2001  CGTCAGGTCC TTTTCAATCG TGTAACCAAA ATAAACTCCA CTAGAAGGAT
      GCAGTCCAGG AAAAGTTAGC ACATTGGTTT TATTTGAGGT GATCTTCCTA

2051  ATTGTGGGGC AACAACACAA TGGGCGTTAC AGGAATATTG CAGTTACCTC
      TAACACCCCG TTGTTGTGTT ACCCGCAATG TCCTTATAAC GTCAATGGAG

2101  GTGATCGATT CAAGAGGACA TCATTCTTTC TTTGGGTAAT TATCCTTTTC
      CACTAGCTAA GTTCTCCTGT AGTAAGAAAG AAACCCATTA ATAGGAAAAG

2151  CAAAGAACAT TTTCCATCCC ACTTGGAGTC ATCCACAATA GCACATTACA
      GTTTCTTGTA AAAGGTAGGG TGAACCTCAG TAGGTGTTAT CGTGTAATGT

2201  GGTTAGTGAT GTCGACAAAC TAGTTTGTCG TGACAAACTG TCATCCACAA
      CCAATCACTA CAGCTGTTTG ATCAAACAGC ACTGTTTGAC AGTAGGTGTT

2251  ATCAATTGAG ATCAGTTGGA CTGAATCTCG AAGGGAATGG AGTGGCAACT
      TAGTTAACTC TAGTCAACCT GACTTAGAGC TTCCCTTACC TCACCGTTGA

2301  GACGTGCCAT CTGCAACTAA AAGATGGGGC TTCAGGTCCG GTGTCCCACC
      CTGCACGGTA GACGTTGATT TTCTACCCCG AAGTCCAGGC CACAGGGTGG

2351  AAAGGTGGTC AATTATGAAG CTGGTGAATG GGCTGAAAAC TGCTACAATC
      TTTCCACCAG TTAATACTTC GACCACTTAC CCGACTTTTG ACGATGTTAG

2401  TTGAAATCAA AAAACCTGAC GGCAGTGAGT GTCTACCAGC AGCGCCAGAC
      AACTTTAGTT TTTTGGACTG CCGTCACTCA CAGATGGTCG TCGCGGTCTG

2451  GCGATTCGGG GCTTCCCCCG GTGCCGGTAT GTGCACAAAG TATCAGGAAC
      CGCTAAGCCC CGAAGGGGGC CACGGCCATA CACGTGTTTC ATAGTCCTTG

2501  GGCACCGTGT GCCGGAGACT TTGCCTTCCA TAAAGAGGGT GCTTTCTTCC
      CCCTGGCACA CGGCCTCTGA AACGGAAGGT ATTTCTCCCA CGAAAGAAGG

2551  TGTATGATCG ACTTGCTTCC ACAGTTATCT ACCGAGGAAC GACTTTCGCT
      ACATACTAGC TGAACGAAGG TGTCAATAGA TGGCTCCTTG CTGAAAGCGA
```

```
2601  GAAGGTGTCG TTGCATTTCT GATACTGCCC CAAGCTAAGA AGGACTTCTT
      CTTCCACAGC AACGTAAAGA CTATGACGGG GTTCGATTCT TCCTGAAGAA

2651  CAGCTCACAC CCCTTGAGAG AGCCGGTCAA TGCAACGGAG GACCCGTCTA
      GTCGAGTGTG GGGAACTCTC TCGGCCAGTT ACGTTGCCTC CTGGGCAGAT

                                  EcoRV
                                 ------

2701  GTGGCTACTA TTCTACCACA ATTAGATATC AGGCTACCGG TTTTGGAACC
      CACCGATGAT AAGATGGTGT TAATCTATAG TCCGATGGCC AAAACCTTGG

2751  AATGACACAG AGTACTTGTT CGAGGTTGAC AATTTGACCT ACGTCCAACT
      TTACTCTGTC TCATGAACAA GCTCCAACTG TTAAACTGGA TGCAGGTTGA

2801  TGAATCAAGA TTCACACCAC AGTTTCTGCT CCAGCTGAAT GAGACAATAT
      ACTTAGTTCT AAGTGTGGTG TCAAAGACGA GGTCGACTTA CTCTGTTATA

2851  ATACAAGTGG GAAAAGGAGC AATACCACGG GAAAACTAAT TTGGAAGGTC
      TATGTTCACC CTTTTCCTCG TTATGGTGCC CTTTTGATTA AACCTTCCAG

2901  AACCCCGAAA TTGATACAAC AATCGGGGAG TGGGCCTTCT GGGAAACTAA
      TTGGGGCTTT AACTATGTTG TTAGCCCCTC ACCCGGAAGA CCCTTTGATT

2951  AAAAACCTCA CTAGAAAAAT TCGCAGTGAA GAGTTGTCTT TCACAGTTGT
      TTTTTGGAGT GATCTTTTTA AGCGTCACTT CTCAACAGAA AGTGTCAACA

3001  ATCAAACGGA GCCAAAAACA TCAGTGGTCA GAGTCCGGCG CGAACTTCTT
      TAGTTTGCCT CGGTTTTTGT AGTCACCAGT CTCAGGCCGC GCTTGAAGAA

3051  CCGACCCAGG GACCAACACA ACAACTGAAG ACCACAAAAT CATGGCTTCA
      GGCTGGGTCC CTGGTTGTGT TGTTGACTTC TGGTGTTTTA GTACCGAAGT

3101  GAAAATTCCT CTGCAATGGT TCAAGTGCAC AGTCAAGGAA GGGAAGCTGC
      CTTTTAAGGA GACGTTACCA AGTTCACGTG TCAGTTCCTT CCCTTCGACG

3151  AGTGTCGCAT CTAACAACCC TTGCCACAAT CTCCACGAGT CCCCAATCCC
      TCACAGCGTA GATTGTTGGG AACGGTGTTA GAGGTGCTCA GGGGTTAGGG

3201  TCACAACCAA ACCAGGTCCG GACAACAGCA CCCATAATAC ACCCGTGTAT
      AGTGTTGGTT TGGTCCAGGC CTGTTGTCGT GGGTATTATG TGGGCACATA

3251  AAACTTGACA TCTCTGAGGC AACTCAAGTT GAACAACATC ACCGCAGAAC
      TTTGAACTGT AGAGACTCCG TTGAGTTCAA CTTGTTGTAG TGGCGTCTTG

3301  AGACAACGAC AGCACAGCCT CCGACACTCC CTCTGCCACG ACCGCAGCCG
      TCTGTTGCTG TCGTGTCGGA GGCTGTGAGG GAGACGGTGC TGGCGTCGGC

3351  GACCCCCAAA AGCAGAGAAC ACCAACACGA GCAAGAGCAC TGACTTCCTG
      CTGGGGGTTT TCGTCTCTTG TGGTTGTGCT CGTTCTCGTG ACTGAAGGAC

3401  GACCCCGCCA CCACAACAAG TCCCCAAAAC CACAGCGACA CCGCTGGCAA
      CTGGGGCGGT GGTGTTGTTC AGGCGTTTTG GTGTCGCTGT GGCGACCGTT

3451  CAACAACACT CATCACCAAG ATACCGGAGA AGAGAGTGCC AGCAGCGGGA
      GTTGTTGTGA GTAGTGGTTC TATGGCCTCT TCTCTCACGG TCGTCGCCCT
```

```
3501  AGCTAGGCTT AATTA:CAAT ACTATTGCTG GAGTCGCAGG ACTGATCACA
      TCGATCCGAA TTAA:.GTTA TGATAACGAC CTCAGCGTCC TGACTAGTGT
.....
3551  GGCGGGAGAA GAACTCGAAG AGAAGCAATT GTCAATGCTC AACCCAAATG
      CCGCCCTCTT CTTGAGCTTC TCTTCGTTAA CAGTTACGAG TTGGGTTTAC
.....
3601  CAACCCTAAT TTACATTACT GGACTACTCA GGATGAAGGT GCTGCAATCG
      GTTGGGATTA AATGTAATGA CCTGATGAGT CCTACTTCCA CGACGTTAGC
.....
3651  GACTGGCCTG GATACCATAT TTCGGGCCAG CAGCCGAGGG AATTTACATA
      CTGACCGGAC CTATGGTATA AAGCCCGGTC GTCGGCTCCC TTAAATGTAT
.....
3701  GAGGGGCTAA TGCACAATCA AGATGGTTTA ATCTGTGGGT TGAGACAGCT
      CTCCCCGATT ACGTGTTAGT TCTACCAAAT TAGACACCCA ACTCTGTCGA
.....
3751  GGCCAACGAG ACGACTCAAG CTCTTCAACT GTTCCTGAGA GCCACAACTG
      CCGGTTGCTC TGCTGAGTTC GAGAAGTTGA CAAGGACTCT CGGTGTTGAC
.....
3801  AGCTACGCAC CTTTTCAATC CTCAACCGTA AGGCAATTGA TTTCTTGCTG
      TCGATGCGTG GAAAAGTTAG GAGTTGGCAT TCCGTTAACT AAAGAACGAC
.....
3851  CAGCGATGGG GCGGCACATG CCACATTCTG GGACCGGACT GCTGTATCGA
      GTCGCTACCC CGCCGTGTAC GGTGTAAGAC CCTGGCCTGA CGACATAGCT
.....
3901  ACCACATGAT TGGACCAAGA ACATAACAGA CAAAATTGAT CAGATTATTC
      TGGTGTACTA ACCTGGTTCT TGTATTGTCT GTTTTAACTA GTCTAATAAG
.....
3951  ATGATTTTGT TGATAAAACC CTTCCGGACC AGGGGGACAA TGACAATTGG
      TACTAAAACA ACTATTTTGG GAAGGCCTGG TCCCCCTGTT ACTGTTAACC
.....
4001  TGGACAGGAT GGAGACAATG GATACCGGCA GGTATTGGAG TTACAGGCGT
      ACCTGTCCTA CCTCTGTTAC CTATGGCCGT CCATAACCTC AATGTCCGCA
.....
4051  TATAATTGCA GTTATCGCTT TATTCTGTAT ATGCAAATTT GTCTTTTAGT
      ATATTAACGT CAATAGCGAA ATAAGACATA TACGTTTAAA CAGAAAATCA
.....
4101  TTTTCTTCAG ATTGCTTCAT GGAAAAGCTC AGCCTCAAAT CAATGAAACC
      AAAAGAAGTC TAACGAAGTA CCTTTTCGAG TCGGAGTTTA GTTACTTTGG
.....
4151  AGGATTTAAT TATATGGATT ACTTGAATCT AAGATTACTT GACAAATGAT
      TCCTAAATTA ATATACCTAA TGAACTTAGA TTCTAATGAA CTGTTTACTA
.....
4201  AATATAATAC ACTGGAGCTT TAAACATAGC CAATGTGATT CTAACTCCTT
      TTATATTATG TGACCTCGAA ATTTGTATCG GTTACACTAA GATTGAGGAA
.....
4251  TAAACTCACA GTTAATCATA AACAAGGTTT GGAATTGATC TGCTGTGCCT
      ATTTGAGTGT CAATTAGTAT TTGTTCCAAA CCTTAACTAG ACGACACGGA
.....
4301  TCTAGTTGCC AGCCATCTGT TGTTTGCCCC TCCCCCGTGC CTTCCTTGAC
      AGATCAACGG TCGGTAGACA ACAAACGGGG AGGGCGCACG GAAGGAACTG
.....
4351  CCTCGAAGGT GCCACTCCCA CTGTCCTTTC CTAATAAAAT GAGGAAATTG
      GGAGCTTCCA CGGTGAGGGT GACAGGAAAG GATTATTTTA CTCCTTTAAC
.....
4401  CATCGCATTG TCTGAGTAGG TGTCATTCTA TTCTGGGGGG TGGGGTGGGG
      GTAGCGTAAC AGACTCATCC ACAGTAAGAT AAGACCCCCC ACCCCACCCC
.....
```

SphI

```
4451  CACCACAGCA AGGGCGAGGA TTGGGAAGAC AATAGCAGGC ATGCTGGGGA
      GTCGTGTCGT TCCCGCTCCT AACCCTTCTG TTATCGTCCG TACGACCCCT
```

KpnI

```
4501  TCCGGTGGGC TCTATGGGTA CCCAGGTGCT GAAGAATTGA CCCGGTTCCT
      AGGCCACCCG AGATACCCAT GGGTCCACGA CTTCTTAACT GGGCCAAGGA
4551  CCTGGGCCAG AAAGAAGCAG GCACATCCCC TTCTCTGTGA CACACCCTGT
      GGACCCGGTC TTTCTTCGTC CGTGTAGGGG AAGAGACACT GTGTGGGACA
4601  CCACGCCCCT GGTTCTTAGT TCCAGCCCCA CTCATAGGAC ACTCATAGCT
      GGTGCGGGGA CCAAGAATCA AGGTCGGGGT GAGTATCCTG TGAGTATCGA
4651  CAGGAGGCCT CCGCCTTCAA TCCCACCCGC TAAAGTACTT GGAGCCGTCT
      GTCCTCCCGA GGCGGAAGTT AGGGTGGGCG ATTTCATGAA CCTCGGCAGA
4701  CTCCCTCCCT CATCAGCCCA CCAAACCAAA CCTAGCCTCC AAGAGTGGGA
      GAGGGAGGGA GTAGTCGGGT GGTTTGGTTT GGATCGGAGG TTCTCACCCT
4751  AGAAACTAAA GCAAGATAGG CTATTAAGTC CAGAGGGAGA GAAAATGCCT
      TCTTTAATTT CGTTCTATCC GATAATTCAG GTCTCCCTCT CTTTTACGGA
4801  CCAACATGTG AGGAAGTAAT GAGAGAAATC ATAGAATTTC TTCCGCTTCC
      GGTTGTACAC TCCTTCATTA CTCTCTTTAG TATCTTAAAG AAGGCGAAGG
4851  TCGCTCACTG ACTCGCTGCG CTCGGTCGTT CGGCTGCGGC GAGCGGTATC
      AGCGAGTGAC TGAGCGACGC GAGCCAGCAA GCCGACGCCG CTCGCCATAG
4901  AGCTCACTCA AAGGCGGTAA TACGGTTATC CACAGAATCA GGGGATAACG
      TCGAGTGAGT TTCCGCCATT ATGCCAATAG GTGTCTTAGT CCCCTATTGC
4951  CAGGAAAGAA CATGTGAGCA AAAGGCCAGC AAAAGGCCAG GAACCGTAAA
      GTCCTTTCTT GTACACTCGT TTTCCGGTCG TTTTCCGGTC CTTGGCATTT
5001  AAGGCCGCGT TGCTGGCGTT TTTCCATAGG CTCCGCCCCC CTGACGAGCA
      TTCCGGCGCA ACGACCCCAA AAAGGTATCC GAGGCGGGGG GACTGCTCGT
5051  TCACAAAAAT CGACGCTCAA GTCAGAGGTG GCGAAACCCG ACAGGACTAT
      AGTGTTTTTA GCTGCGAGTT CAGTCTCCAC CGCTTTGGGC TGTCCTGATA
5101  AAAGATACCA GGCGTTTCCC CCTGGAAGCT CCCTCGTGCG CTCTCCTGTT
      TTTCTATGGT CCGCAAAGGG GGACCTTCGA GGGAGCACGC GAGAGGACAA
5151  CCGACCCTGC CGCTTACCGG ATACCTGTCC GCCTTTCTCC CTTCGGGAAG
      GGCTGGGACG GCGAATGGCC TATGGACAGG CGGAAAGAGG GAAGCCCTTC
5201  CGTGGCGCTT TCTCAATGCT CACGCTGTAG GTATCTCAGT TCGGTGTAGG
      GCACCGCGAA AGAGTTACGA GTGCGACATC CATAGAGTCA AGCCACATCC
5251  TCGTTCGCTC CAAGCTGGGC TGTGTGCACG AACCCCCCGT TCAGCCCGAC
      AGCAAGCGAG GTTCGACCCG ACACACGTGC TTGGGGGGCA AGTCGGGCTG
5301  CGCTGCGCCT TATCCGGTAA CTATCGTCTT GAGTCCAACC CGGTAAGACA
      GCGACGCGGA ATAGGCCATT GATAGCAGAA CTCAGGTTGG GCCATTCTGT
```

```
5351  CGACTTATCG CCACTGGCAG CAGCCACTGG TAACAGGATT AGCAGAGCGA
      GCTGAATAGC GGTGACCGTC GTCGGTGACC ATTGTCCTAA TCGTCTCGCT

5401  GGTATGTAGG CGGTGCTACA GAGTTCTTGA AGTGGTGGCC TAACTACGGC
      CCATACATCC GCCACGATGT CTCAAGAACT TCACCACCGG ATTGATGCCG

5451  TACACTAGAA GGACAGTATT TGGTATCTGC GCTCTGCTGA AGCCAGTTAC
      ATGTGATCTT CCTGTCATAA ACCATAGACG CGAGACGACT TCGGTCAATG

5501  CTTCGGAAAA AGAGTTGGTA GCTCTTGATC CGGCAAACAA ACCACCGCTG
      GAAGCCTTTT TCTCAACCAT CGAGAACTAG GCCGTTTGTT TGGTGGCGAC

5551  GTAGCGGTGG TTTTTTTGTT TGCAAGCAGC AGATTACGCG CAGAAAAAAA
      CATCGCCACC AAAAAAACAA ACGTTCGTCG TCTAATGCGC GTCTTTTTTT

5601  GGATCTCAAG AAGATCCTTT GATCTTTTCT ACGGGGTCTG ACGCTCAGTG
      CCTAGAGTTC TTCTAGGAAA CTAGAAAAGA TGCCCCAGAC TGCGAGTCAC

5651  GAACGAAAAC TCACGTTAAG GGATTTTGGT CATGAGATTA TCAAAAAGGA
      CTTGCTTTTG AGTGCAATTC CCTAAAACCA GTACTCTAAT AGTTTTTCCT

5701  TCTTCACCTA GATCCTTTTA AATTAAAAAT GAAGTTTTAA ATCAATCTAA
      AGAAGTGGAT CTAGGAAAAT TTAATTTTTA CTTCAAAATT TAGTTAGATT

5751  AGTATATATG AGTAAACTTG GTCTGACAGT TACCAATGCT TAATCAGTGA
      TCATATATAC TCATTTGAAC CAGACTGTCA ATGGTTACGA ATTAGTCACT

5801  GGCACCTATC TCAGCGATCT GTCTATTTCG TTCATCCATA GTTGCCTGAC
      CCGTGGATAG AGTCGCTAGA CAGATAAAGC AAGTAGGTAT CAACGGACTG

5851  TCCCCGGGGG GGGGGCGCTG AGGTCTGCCT CGTGAAGAAG GTGTTGCTGA
      AGGGGCCCCC CCCCCGCGAC TCCAGACGGA GCACTTCTTC CACAACGACT

5901  CTCATACCAG GCCTGAATCG CCCCATCATC CAGCCAGAAA GTGAGGGAGC
      GAGTATGGTC CGGACTTAGC GGGGTAGTAG GTCGGTCTTT CACTCCCTCG

5951  CACGGTTGAT GAGAGCTTTG TTGTAGGTGG ACCAGTTGGT GATTTTGAAC
      GTGCCAACTA CTCTCGAAAC AACATCCACC TGGTCAACCA CTAAAACTTG

6001  TTTTGCCTTG CCACGGAACG GTCTGCGTTG TCGGGAAGAT GCGTGATCTG
      AAAACGAAAC GGTGCCTTGC CAGACGCAAC AGCCCTTCTA CGCACTAGAC

6051  ATCCTTCAAC TCAGCAAAAG TTCGATTTAT TCAACAAAGC CGCCGTCCCG
      TAGGAAGTTG AGTCGTTTTC AAGCTAAATA AGTTGTTTCG GCGGCAGGGC

6101  TCAAGTCAGC GTAATGCTCT GCCAGTGTTA CAACCAATTA ACCAATTCTG
      AGTTCAGTCG CATTACGAGA CGGTCACAAT GTTGGTTAAT TGGTTAAGAC

6151  ATTAGAAAAA CTCATCGAGC ATCAAATGAA ACTGCAATTT ATTCATATCA
      TAATCTTTTT GAGTAGCTCG TAGTTTACTT TGACGTTAAA TAAGTATAGT

6201  GGATTATCAA TACCATATTT TTGAAAAAGC CGTTTCTGTA ATGAAGGAGA
      CCTAATAGTT ATGGTATAAA AACTTTTTCG GCAAAGACAT TACTTCCTCT

6251  AAACTCACCG AGGCAGTTCC ATAGGATGGC AAGATCCTGG TATCGGTCTG
      TTTGAGTGGC TCCGTCAAGG TATCCTACCG TTCTAGGACC ATAGCCAGAC
```

```
6301  CGATTCCGAC TCGTCCAACA TCAATACAAC CTATTAATTT CCCCTCGTCA
      GCTAAGGCTG AGCAGGTTGT AGTTATGTTG GATAATTAAA GGGGAGCAGT
.....
6351  AAAATAAGGT TATCAAGTGA GAAATCACCA TGAGTCACGA CTCAATCCGG
      TTTTATTCCA ATAGTTCACT CTTTAGTGGT ACTCACTGCT GACTTAGGCC
.....
                 HindIII
                 ========
6401  TGAGAATGGC AAAAGCTTAT GCATTTCTTT CCAGACTTGT TCAACAGGCC
      ACTCTTACCG TTTTCGAATA CGTAAAGAAA GGTCTGAACA AGTTGTCCGG
.....
6451  AGCCATTACG CTCGTCATCA AAATCACTCG CATCAACCAA ACCGTTATTC
      TCGGTAATGC GAGCAGTAGT TTTAGTGAGC GTAGTTGGTT TGGCAATAAG
.....
                                        PvuI
                                        ====
6501  ATTCGTGATT GCGCCTGAGC GAGACGAAAT ACGCGATCGC TGTTAAAAGG
      TAAGCACTAA CGCGGACTCG CTCTGCTTTA TGCGCTAGCG ACAATTTCC
.....
6551  ACAATTACAA ACAGGAATCG AATGCAACCG GCGCAGGAAC ACTGCCAGCG
      TGTTAATGTT TGTCCTTAGC TTACGTTGGC CGCGTCCTTG TGACGGTCGC
.....
6601  CATCAACAAT ATTTTCACCT GAATCAGGAT ATTCTTCTAA TACCTGGAAT
      GTAGTTGTTA TAAAAGTGGA CTTAGTCCTA TAAGAAGATT ATGGACCTTA
.....
6651  GCTGTTTCCC CGGGGATCGC AGTGGTGAGT AACCATGCAT CATCAGGAGT
      CGACAAAAGG GCCCCTAGCG TCACCACTCA TTGGTACGTA GTAGTCCTCA
.....
6701  ACGGATAAAA TGCTTGATGG TCGGAAGAGG CATAAATTCC GTCAGCCAGT
      TGCCTATTTT ACGAACTACC AGCCTTCTCC GTATTTAAGG CAGTCGGTCA
.....
6751  TTACTCTGAC CATCTCATCT GTAACATCAT TGGCAACGCT ACCTTTGCCA
      AATGAGACTG GTAGAGTAGA CATTGTAGTA ACCGTTGCGA TGGAAACGGT
.....
6801  TGTTTCAGAA ACAACTCTGG CGCATCGGGC TTCCCATACA ATCGATAGAT
      ACAAAGTCTT TGTTGAGACC GCGTAGCCCG AAGGGTATGT TAGCTATCTA
.....
6851  TGTCGCACCT GATTGCCCGA CATTATCGCG AGCCCATTTA TACCCATATA
      ACAGCGTGGA CTAACGGGCT GTAATAGCGC TCGGGTAAAT ATGGGTATAT
.....
                                        XhoI
                                        ====
6901  AATCAGCATC CATGTTGGAA TTTAATCGCG GCCTCGAGCA AGACGTTTCC
      TTAGTCGTAG GTACAACCTT AAATTAGCGC CGGAGCTCGT TCTGCAAAGG
.....
6951  CGTTGAATAT GGCTCATAAC ACCCCTTGTA TTACTGTTTA TGTAAGCAGA
      GCAACTTATA CCGAGTATTG TGGGGAACAT AATGACAAAT ACATTCGTCT
.....
7001  CAGTTTTATT GTTCATGATG ATATATTTTT ATCTTGTGCA ATGTAACATC
      GTCAAAATAA CAAGTACTAC TATATAAAAA TAGAACACGT TACATTGTAG
.....
                 DraIII
                 ======
7051  AGAGATTTTG AGACACAACG TGGCTTTCCC CCCCCCCCCA TTATTGAAGC
      TCTCTAAAAC TCTGTGTTGC ACCGAAAGGG GGGGGGGGGT AATAACTTCG
.....
7101  ATTTATCAGG GTTATTGTCT CATGAGCGGA TACATATTTG AATGTATTTA
      TAAATAGTCC CAATAACAGA GTACTCGCCT ATGTATAAAC TTACATAAAT
.....
```

```
7151  GAAAAATAAA CAAATAGGGG TTCCGCGCAC ATTTCCCCGA AAAGTGCCAC
      CTTTTTATTT GTTTATCCCC AAGGCGCGTG TAAAGGGGCT TTTCACGGTG
.......................................................................
7201  CTGACGTCTA AGAAACCATT ATTATCATGA CATTAACCTA TAAAAATAGG
      GACTGCAGAT TCTTTGGTAA TAATAGTACT GTAATTGGAT ATTTTTATCC
.......................................................................
7251  CGTATCACGA GGCCCTTTCG TC
      GCATAGTGCT CCGGGAAAGC AG
.......................................................................
```